

**IDENTIFIKASI KEBUTUHAN DAN PREFERENSI
KONSUMEN TERHADAP PRODUK *HEADSET*
BERDASARKAN *ONLINE REVIEWS***

SKRIPSI

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

Disusun oleh:

Nama : Yoko Gunawan Candra

NPM : 2017610129



**PROGRAM STUDI SARJANA TEKNIK INDUSTRI
JURUSAN TEKNIK INDUSTRI
FAKULTAS TEKNOLOGI INDUSTRI
UNIVERSITAS KATOLIK PARAHYANGAN
2021**

**IDENTIFIKASI KEBUTUHAN DAN PREFERENSI
KONSUMEN TERHADAP PRODUK *HEADSET*
BERDASARKAN *ONLINE REVIEWS***

SKRIPSI

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

Disusun oleh:

Nama : Yoko Gunawan Candra

NPM : 2017610129



**PROGRAM STUDI SARJANA TEKNIK INDUSTRI
JURUSAN TEKNIK INDUSTRI
FAKULTAS TEKNOLOGI INDUSTRI
UNIVERSITAS KATOLIK PARAHYANGAN
2021**

**FAKULTAS TEKNOLOGI INDUSTRI
UNIVERSITAS KATOLIK PARAHYANGAN
BANDUNG**



Nama : Yoko Gunawan Candra
NPM : 2017610129
Jurusan : Teknik Industri
Judul Skripsi : IDENTIFIKASI KEBUTUHAN DAN PREFERENSI KONSUMEN
TERHADAP PRODUK *HEADSET* BERDASARKAN *ONLINE
REVIEWS*

TANDA PERSETUJUAN SKRIPSI

Bandung, 1 Maret 2021

Ketua Program Studi Sarjana Teknik Industri

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

Pembimbing Pertama

4 Maret 2021

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



Program Studi Sarjana Teknik Industri
Jurusan Teknik Industri
Fakultas Teknologi Industri
Universitas Katolik Parahyangan

Pernyataan Tidak Mencontek atau Melakukan Tindakan Plagiat

Saya, yang bertanda tangan di bawah ini,

Nama : Yoko Gunawan Candra

NPM : 2017610129

dengan ini menyatakan bahwa skripsi dengan judul :

“IDENTIFIKASI KEBUTUHAN DAN PREFERENSI KONSUMEN TERHADAP PRODUK *HEADSET* BERDASARKAN *ONLINE REVIEWS*”

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, 1 Maret 2021

Yoko Gunawan Candra
2017610129

ABSTRAK

Headset merupakan produk yang telah banyak digunakan dalam kehidupan sehari-hari dan semakin dibutuhkan dalam kegiatan daring pada saat pandemi COVID-19. Untuk bersaing di pasar *headset* yang terus meningkat, keinginan pasar harus dipenuhi. Untuk memenuhi keinginan pasar, kebutuhan konsumen perlu dipahami. Selain kebutuhan konsumen, preferensi konsumen juga penting dalam menjelaskan perilaku konsumen. Seiring berjalannya waktu, penggunaan internet terus meningkat, begitu juga data *online reviews* produk *headset* yang dapat dimanfaatkan untuk mengetahui kebutuhan dan preferensi konsumen. Hal ini juga menimbulkan tantangan bagi perusahaan yaitu data yang kurang terstruktur dan waktu pengolahan data yang lebih lama.

Latent Dirichlet Allocation (LDA) dan *Non-Negative Matrix Factorization* (NMF) merupakan metode *topic modelling* yang dapat digunakan untuk menentukan aspek-aspek produk berdasarkan data *review*. Setelah itu, dapat dilakukan analisis sentimen untuk membandingkan sentimen terhadap setiap aspek produk. Melalui aspek-aspek yang teridentifikasi, juga dapat diidentifikasi kebutuhan dan preferensi konsumen. LDA merupakan metode probabilistik dan NMF merupakan metode non-probabilistik. Pada penelitian ini, akan dilakukan perbandingan *topic modelling* dengan *Latent Dirichlet Allocation* (LDA) dan *Non-Negative Matrix Factorization* (NMF).

Berdasarkan pengolahan data, aspek-aspek yang teridentifikasi adalah durabilitas produk, penggunaan produk, *noise cancellation*, bagian-bagian produk, kualitas produk secara umum, dan *customer service*. Kebutuhan dan preferensi konsumen diidentifikasi berdasarkan aspek-aspek tersebut. Analisis sentimen pada *headset* yang mengalami penurunan *rating* menunjukkan perubahan skor sentimen untuk setiap aspek sebelum dan sesudah tahun 2020. Usulan diberikan untuk produk yang sudah ada di pasaran dan untuk produk yang sedang dalam pengembangan berdasarkan kebutuhan dan preferensi yang teridentifikasi. Berdasarkan perbandingan metode LDA dan NMF, LDA lebih baik karena terdapat lebih sedikit kata kunci yang muncul bersamaan dalam beberapa aspek sekaligus dibandingkan dengan NMF.

ABSTRACT

Headsets are a product that has been widely used in everyday life and is increasingly needed in online activities during the COVID-19 pandemic. In order to compete in the increasing headset market, market wants must be fulfilled. To fulfill market wants, consumer needs need to be realized. Apart from consumer needs, consumer preferences are also important in explaining consumer behavior. Over time, internet usage continues to increase, as well as headset product's online reviews data that can be used to understand consumer needs and preferences. This also poses challenges for companies, namely unstructured data and longer data processing time.

Latent Dirichlet Allocation (LDA) dan Non-Negative Matrix Factorization (NMF) are topic modelling methods that can be used to determine product aspects based on the review data. Afterwards, sentiment analysis can be done to compare the sentiment for each aspect of the product. Through the identified aspects, consumer needs and preferences can also be identified. LDA is a probabilistic method and NMF is a non-probabilistic method. This research will compare the topic modelling using Latent Dirichlet Allocation (LDA) and Non-Negative Matrix Factorization.

Based on the data processing, the identified aspects are product durability, product use, noise cancellation, product parts, product quality in general, and customer service. Consumer needs and preferences are identified based on these aspects. Sentiment analysis on headset that have decreased rating shows changes in sentiment scores for each aspect before and after 2020. Suggestions are made for products that have been on the market and for products under development based on identified needs and preferences. Based on the comparison of LDA and NMF method, LDA performs better because there are fewer concurrent keywords in several aspects at the same time compared to NMF.

KATA PENGANTAR

Puji dan Syukur penulis panjatkan kepada Tuhan Yang Maha Esa karena atas berkat dan kasih-Nya laporan skripsi dengan judul “Identifikasi Kebutuhan dan Preferensi Konsumen Terhadap Produk *Headset* Berdasarkan *Online Reviews*” dapat diselesaikan. Laporan ini dibuat sebagai salah satu syarat kelulusan untuk mendapatkan gelar Sarjana Teknik Industri di Universitas Katolik Parahyangan. Laporan ini dapat diselesaikan berkat dukungan, bantuan, serta bimbingan dari berbagai pihak. Penulis ingin mengucapkan terima kasih sebanyak-banyaknya kepada pihak-pihak yang telah membantu dan mendukung penulis dalam mengerjakan skripsi ini. Berikut ucapan terima kasih kepada pihak-pihak yang telah mendukung penulis.

1. Bapak Dedy Suryadi, S.T., M.S., Ph.D. selaku dosen pembimbing skripsi yang telah memimbing penulis dalam mengerjakan penelitian ini.
2. Keluarga penulis, khususnya orangtua dan saudara penulis yang selalu mendukung dan memberi semangat dalam mengerjakan penulisan laporan ini.
3. Teman-teman kelas A khususnya Jonathan Alvin, Nicholas Putra, Arvin Jonathan, Irfan Widiyanto, Patrick Owen, Felita, dan teman-teman lain yang tidak dapat disebutkan satu per satu yang senantiasa memberi dukungan serta semangat dalam mengerjakan skripsi ini.
4. Teman-teman “*All in One Team*” yang menjadi kelompok praktikum dari Proses Manufaktur hingga Otomasi yang terdiri dari Felicia Nathania, Jonathan, Ricky Vincentius, dan Stanley Surya yang selalu mendukung pengerjaan skripsi.
5. Teman-teman kelompok PST “*Solver-V*” yang terdiri dari Hendry Gunawan, Ignatius Kevin, Nicholas Shaul, dan Tejakusuma yang turut mendukung dan memberi semangat selama pengerjaan skripsi.
6. Teman-teman komsel dan gereja penulis, khususnya AOG 23 yang selalu memberi doa dan dukungan untuk pengerjaan skripsi ini.

7. Teman-teman Teknik Industri angkatan 2017 yang tidak dapat disebutkan satu per satu atas kebersamaan dan kerjasamanya selama masa-masa perkuliahan.
8. Pihak-pihak lain yang berkontribusi dalam skripsi ini dan tidak dapat disebutkan satu per satu.

Terima kasih penulis ucapkan kepada pihak-pihak tersebut atas dukungan, semangat, dan doa yang diberikan kepada penulis. Tanpa dukungan dari pihak-pihak tersebut, laporan ini belum tentu dapat diselesaikan dengan baik. Penulis menyadari, masih terdapat berbagai kekurangan dalam penulisan skripsi ini. Penulis meminta maaf atas segala kekurangan maupun kesalahan pada penulisan laporan skripsi ini. Penulis selalu berharap laporan ini dapat berguna bagi pembaca dan menjadi penyemangat untuk melakukan penelitian-penelitian yang berhubungan. Sekali lagi, penulis ucapkan terima kasih banyak.

Bandung, Maret 2021

Yoko Gunawan Candra

DAFTAR ISI

ABSTRAK	i
<i>ABSTRACT</i>	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 Rumusan Masalah	I-5
I.3 Pembatasan Masalah dan Asumsi Penelitian	I-9
I.4 Tujuan Penelitian	I-10
I.5 Manfaat Penelitian	I-10
I.6 Metodologi Penelitian.....	I-10
I.7 Sistematika Laporan	I-15
BAB II TINJAUAN PUSTAKA	II-1
II.1 Kebutuhan dan Preferensi Konsumen.....	II-1
II.2 <i>Data Mining</i> dan <i>Text Mining</i>	II-2
II.3 <i>Sentiment Analysis</i>	II-3
II.3.1 <i>Lexicon</i>	II-4
II.4 <i>Topic Modelling</i>	II-4
II.4.1 <i>Latent Dirichlet Allocation</i>	II-4
II.4.2 <i>Non-Negative Matrix Factorization</i>	II-6
BAB III PENGUMPULAN DAN PENGOLAHAN DATA	III-1
III.1 Pengumpulan Data	III-1
III.2 <i>Topic Modelling</i> dengan <i>Latent Dirichlet Allocation</i>	III-2
III.3 Analisis Sentimen untuk Setiap Aspek.....	III-7
III.3.1 Penentuan Topik untuk Setiap Kalimat	III-7
III.3.2 Analisis Sentimen setiap Kalimat	III-9

III.4 Identifikasi Perubahan Preferensi Konsumen.....	III-11
III.5 <i>Topic Modelling</i> dengan <i>Non-Negative Matrix Factorization</i>	III-16
III.6 Perbandingan Kebutuhan Konsumen berdasarkan LDA dan NMF	III-18
III.7 Pemberian Usulan	III-20
BAB IV ANALISIS	IV-1
IV.1 Analisis Pengumpulan Data.....	IV-1
IV.2 Analisis Pra-Pemrosesan Data	IV-3
IV.3 Analisis Aspek-Aspek yang Teridentifikasi.....	IV-5
IV.4 Analisis Kelemahan Metode	IV-6
IV.5 Analisis Kebutuhan dan Preferensi Konsumen	IV-8
IV.6 Analisis Perbedaan LDA dan NMF	IV-9
IV.7 Analisis Usulan.....	IV-10
BAB V KESIMPULAN SARAN	V-1
V.1 Kesimpulan.....	V-1
V.2 Saran.....	V-2

DAFTAR PUSTAKA

LAMPIRAN

DAFTAR RIWAYAT HIDUP

DAFTAR TABEL

Tabel I.1	Perbandingan Pendekatan Identifikasi Kebutuhan Konsumen.....	I-2
Tabel III.1	Contoh Data yang Dikumpulkan	III-2
Tabel III.2	Hasil LDA	III-6
Tabel III.3	Contoh Pemecahan Komentar.....	III-8
Tabel III.4	Contoh Hasil Analisis Sentimen per Kalimat	III-10
Tabel III.5	Rekapitulasi Sentimen Aspek 0	III-11
Tabel III.6	Hasil NMF.....	III-17
Tabel III.7	Rekapitulasi Hasil LDA dan NMF	III-19
Tabel III.8	Perbandingan Skor Sentimen Produk 12	III-20
Tabel III.9	Interpretasi Kalimat Produk 12.....	III-21
Tabel III.10	Contoh Interpretasi Kebutuhan Konsumen	III-23
Tabel III.11	Kebutuhan Konsumen Terinterpretasi.....	III-24
Tabel IV.1	Hasil LDA 4 Topik.....	IV-6
Tabel IV.2	Proporsi Kata-Kata Penggunaan tiap Produk.....	IV-11

DAFTAR GAMBAR

Gambar I.1	Estimasi Jumlah Pengguna Internet dari Tahun 2005-2009	I-3
Gambar I.2	Tampilan <i>Review</i> di <i>Website Amazon</i>	I-7
Gambar I.3	Grafik <i>Rating</i> Produk <i>Headset X</i>	I-8
Gambar I.4	<i>Flow Chart</i> Metodologi Penelitian	I-11
Gambar II.1	<i>Plate Notation</i> LDA.....	II-5
Gambar II.2	Matriks <i>V</i>	II-7
Gambar II.3	Matriks <i>W</i>	II-7
Gambar II.4	Matriks <i>H</i>	II-7
Gambar III.1	<i>Dependency Parsing</i>	III-4
Gambar III.2	Penentuan Topik Kalimat	III-9
Gambar III.3	Proporsi Frekuensi Kalimat Aspek 0.....	III-12
Gambar III.4	Proporsi Frekuensi Kalimat Aspek 1.....	III-13
Gambar III.5	Proporsi Frekuensi Kalimat Aspek 2.....	III-13
Gambar III.6	Proporsi Frekuensi Kalimat Aspek 3.....	III-14
Gambar III.7	Proporsi Frekuensi Kalimat Aspek 4.....	III-15
Gambar III.8	Perbandingan Proporsi Kalimat Aspek 0,1, dan 2	III-16
Gambar IV.1	<i>Headset</i> yang Digunakan	IV-2

DAFTAR LAMPIRAN

LAMPIRAN A REKAPITULASI <i>COHERENCE SCORE</i>	A-1
LAMPIRAN B REKAPITULASI SENTIMEN SETIAP ASPEK	B-1
LAMPIRAN C INTERPRETASI KEBUTUHAN KONSUMEN	C-1
LAMPIRAN D INTERPRETASI KOMENTAR PRODUK 12	D-1

BAB I

PENDAHULUAN

Bab ini berisi pendahuluan yang dilakukan dalam penelitian. Pendahuluan yang dimulai melalui latar belakang masalah. Setelah itu, dilanjutkan dengan identifikasi rumusan masalah, penentuan tujuan dan manfaat penelitian, serta pemaparan metodologi penelitian.

I.1 Latar Belakang Masalah

Headset merupakan produk yang telah digunakan dalam kehidupan sehari-hari. Pada saat penelitian ini dimulai, peran *headset* semakin dibutuhkan karena adanya pandemi COVID-19. Pandemi COVID-19 menyebabkan berbagai aktivitas yang awalnya dilakukan secara luring menjadi dilakukan secara daring. Peran *headset* yang dimaksud adalah untuk menunjang keperluan pertemuan dan komunikasi *online* seperti *work from home*, *study from home*, *webinar*, dan lain-lain.

Pasar dari *headset* mengungkapkan perkembangan yang cepat pada tahun 2019 serta beberapa tahun terakhir dan akan terus berkembang dengan perkembangan berkelanjutan pada tahun-tahun mendatang. Faktor utama yang mendukung hal tersebut termasuk peningkatan adopsi pada komunikasi, edukasi, hiburan, dan lain-lain di seluruh dunia (Pune, 2019). Pemasaran produk *headset* saat ini telah memanfaatkan perkembangan teknologi informasi, salah satunya adalah melalui jual beli di situs-situs internet. Berdasarkan laporan yang diterbitkan oleh *Marketsandresearch.biz* (2020), pasar *headset* secara global pada tahun 2020-2025 diperkirakan akan terus bertumbuh. Berdasarkan hal tersebut, produk *headset* memiliki peluang pasar yang relatif baik di masa depan. Ukuran pasar sendiri merupakan sinyal dari adanya permintaan. Permintaan yaitu keinginan akan suatu produk spesifik yang didukung oleh daya beli. Hal ini berarti permintaan untuk produk *headset* diproyeksi akan terus meningkat.

Agar dapat bersaing di pasar, suatu perusahaan harus dapat berinovasi mengembangkan produk dan harus dapat memasarkan produknya dengan baik agar dapat menyesuaikan dengan keinginan pasar. Keinginan yaitu kebutuhan

yang mengarah pada objek spesifik. Agar dapat menyesuaikan dengan keinginan pasar, perlu diketahui kebutuhan konsumen. Kebutuhan konsumen adalah hal-hal yang diperlukan oleh konsumen saat membeli barang atau jasa.

Proses identifikasi kebutuhan konsumen sendiri merupakan salah satu tahap dalam pengembangan produk. Tujuan dari tahap ini adalah untuk memahami kebutuhan konsumen dan mengkomunikasikannya kepada tim pengembangan produk (Ulrich & Eppinger, 2012). Dalam proses identifikasi kebutuhan konsumen, kebutuhan konsumen bisa didapatkan melalui berbagai sumber seperti wawancara, diskusi grup, maupun melalui *feedback* dalam bentuk *review* dari konsumen terhadap produk yang sudah ada. Seiring perkembangan teknologi, *review* dari konsumen dapat diberikan melalui berbagai macam media, salah satunya adalah internet. Masing-masing pendekatan tersebut memiliki kelebihan dan kekurangan tersendiri. Tabel I.1 menunjukkan kelebihan dan kekurangan untuk masing-masing pendekatan untuk mengidentifikasi kebutuhan konsumen (Timoshenko, 2019; Ulrich & Eppinger, 2012).

Tabel I.1 Perbandingan Pendekatan Identifikasi Kebutuhan Konsumen

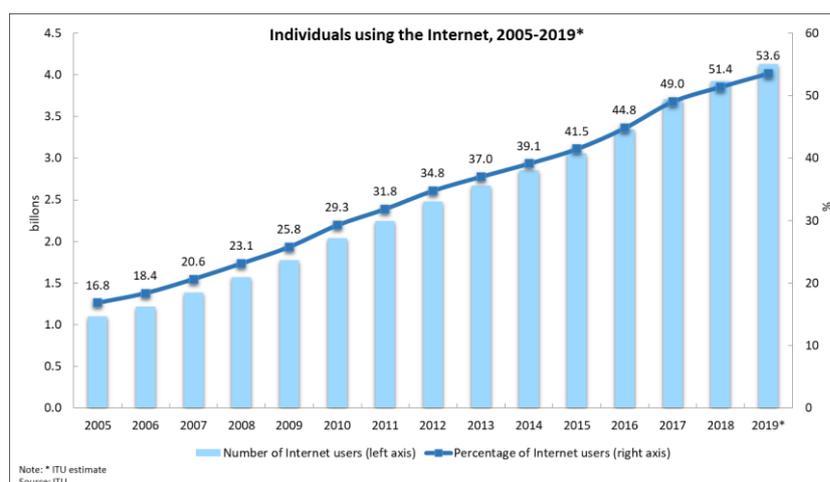
Metode/ Pendekatan	Kelebihan	Kekurangan
Wawancara	Memungkinkan tim pengembangan produk untuk mengalami lingkungan penggunaan produk.	Biaya tinggi
		Membutuhkan waktu lama
Diskusi grup	Interaksi di antara peserta diskusi memungkinkan munculnya kebutuhan yang lebih bervariasi	Biaya tinggi
		Membutuhkan waktu lama
<i>Online Reviews</i>	Dapat memberi akses kepada konsumen yang sulit untuk berpartisipasi dalam wawancara atau diskusi grup	Pada umumnya mengarah ke sangat positif atau sangat negatif, sehingga kebutuhan konsumen yang biasanya diekspresikan dalam bahasa netral bisa terlewat
	<i>Reviewer</i> cenderung langsung menulis <i>review</i> setelah menggunakan produk	

Berdasarkan Tabel I.1, setiap pendekatan memiliki kelebihan dan kekurangan masing-masing. Berdasarkan kelebihan dan kekurangan yang ada, *online review* dapat menjadi pendekatan yang baik dalam mengidentifikasi kebutuhan konsumen. Pada *online review*, *reviewer* cenderung langsung menulis *review* setelah menggunakan produk. Dengan demikian, saat melihat *online*

review secara tidak langsung pengalaman pengguna produk dapat diketahui. Hal lain yang menjadi pertimbangan adalah data *review* dari konsumen terhadap produk sudah tersedia secara *online* sehingga tidak perlu mengumpulkan data lagi. Dengan demikian, waktu dan biaya yang dibutuhkan dapat dikurangi.

Untuk produk *headset*, *online review* bisa dimanfaatkan untuk mengetahui kekuatan dan kelemahan *headset*. Selain itu, *online review* juga bisa dimanfaatkan sebagai pertimbangan dalam mengembangkan *headset* berdasarkan pengalaman pengguna *headset* yang sudah ada di pasaran. Untuk memperlengkap kebutuhan konsumen dalam menjelaskan perilaku konsumen, terdapat preferensi konsumen. Preferensi konsumen adalah keputusan konsumen untuk menyukai atau tidak menyukai sebuah pilihan berdasarkan penilaiannya terhadap produk (Mowen & Minor dalam Albari & Amalia, 2008). Dalam membentuk preferensi, digunakan asumsi bahwa konsumen melihat produk sebagai sekumpulan atribut. Oleh karena itu, preferensi konsumen mengacu pada karakteristik tertentu yang diinginkan oleh konsumen dalam suatu barang atau jasa untuk membuat barang atau jasa tersebut semakin disukai.

Seiring berjalannya waktu, penggunaan internet terus mengalami peningkatan. Hal ini juga dipengaruhi oleh semakin mudahnya akses internet dan juga didukung oleh perkembangan teknologi yang kian pesat. Gambar I.1 adalah grafik yang menunjukkan estimasi jumlah pengguna *internet* dari tahun ke tahun.



Gambar I.1 Estimasi Jumlah Pengguna Internet dari Tahun 2005-2019
(Sumber : <https://www.itu.int/en/ITU-D/Statistics/Pages/stat/default.aspx>)

Gambar I.1 diambil dari *website International Telecommunication Union*. Berdasarkan gambar I.1, pengguna internet pada tahun 2005 diestimasi sekitar 1,1 miliar pengguna atau sekitar 16% dari populasi dunia dan terus mengalami peningkatan. Pada tahun 2019, estimasi pengguna internet mencapai sekitar 4,1 miliar atau setara dengan 53,6% dari populasi dunia. Berdasarkan data tersebut, jumlah pengguna internet memiliki tren untuk terus meningkat. Per Juli 2020, jumlah pengguna internet di seluruh dunia diestimasi sekitar 4,57 miliar pengguna (Clement, 2020).

Peningkatan jumlah pengguna *internet* membarengi peningkatan jumlah data yang beredar di internet, tak terkecuali data *review* terhadap produk. *Review* produk merupakan ulasan pengguna mengenai kualitas suatu produk atau jasa. Dalam konteks *e-commerce*, *review* produk digunakan oleh situs perbelanjaan untuk memberikan kesempatan kepada pelanggan untuk menilai dan mengomentari produk yang telah mereka beli. Dengan demikian, suatu *review* produk berisi penilaian yang pada umumnya dalam bentuk *rating* dan komentar terhadap suatu produk.

Di dalam komentar terhadap produk, terdapat sentimen dan opini yang memiliki karakteristik subjektif. Karakteristik subjektif pada sentimen dan opini disebabkan oleh perbedaan pada pengalaman, cara pandang, dan ketertarikan masing-masing individu (Liu, 2015). Karena banyaknya perbedaan tersebut, penting untuk dapat memperhatikan opini dari banyak orang ketimbang opini dari satu orang saja. Opini dari satu orang hanya akan merepresentasikan pengalaman, cara pandang, dan ketertarikan dari satu orang saja sehingga menjadi kurang relevan untuk produk yang dipakai oleh banyak orang. Opini penting bagi perusahaan karena opini dapat menunjukkan pandangan dari konsumen mengenai produknya. Oleh karena itu, perlu dibentuk suatu kesimpulan tertentu yang dapat merepresentasikan opini banyak orang.

Berdasarkan hal tersebut, peningkatan jumlah data *review* yang beredar di internet dapat menjadi peluang bagi perusahaan untuk mengidentifikasi kebutuhan dan preferensi konsumen. Kebutuhan konsumen dapat dimanfaatkan untuk mengembangkan produknya agar sesuai dengan keinginan konsumen. Dengan membaca keinginan konsumen melalui sentimen dan opini dalam komentar, produk yang dikembangkan dapat lebih menyesuaikan dengan keinginan pasar. Sedangkan preferensi konsumen dapat dimanfaatkan untuk

meningkatkan daya saing produk. Dengan demikian, produk dapat semakin laku di pasaran dan meningkatkan pemasukan bagi perusahaan tersebut.

Selain peluang, banyaknya jumlah data juga menimbulkan tantangan tersendiri bagi perusahaan. Tantangan yang timbul adalah untuk mengolah data-data yang ada. Salah satu tantangan yang muncul adalah waktu pengolahan data yang lebih lama karena jumlah data yang lebih banyak. Waktu pengolahan data yang lama dapat diatasi dengan penggunaan komputer. Namun, penggunaan komputer untuk mengolah data *review* juga memiliki tantangan tersendiri. Komentar pada umumnya berbentuk data tekstual yang cenderung lebih tidak terstruktur dibandingkan dengan data kuantitatif. Oleh karena itu, data tekstual perlu diolah terlebih dahulu sebelum dapat dianalisa lebih lanjut dengan bantuan komputer. Untuk memanfaatkan peluang dan menghadapi tantangan yang ada di balik semakin meningkatnya jumlah data, perlu dirancang suatu sistem yang dapat mempermudah proses tersebut.

I.2 Identifikasi Rumusan Masalah

Pada subbab latar belakang masalah telah dijelaskan bahwa produk *headset* memiliki peluang pasar yang relatif baik di masa depan. Namun, pasar *headset* juga memiliki tantangan tersendiri. Tantangan tersebut antara lain persaingan pasar dengan produk-produk lain, perubahan teknologi yang cepat, serta cepatnya perubahan keinginan pasar. Untuk memanfaatkan peluang dan mengatasi tantangan tersebut, perusahaan perlu memahami kebutuhan dan preferensi konsumen terhadap produk *headset*. Namun, kebutuhan dan preferensi konsumen untuk produk *headset* tidak selalu sama. Sebagai *consumer goods* atau *final goods*, produk *headset* terus mengalami perubahan kebutuhan dan preferensi konsumen. Hal tersebut menjadi tantangan dalam mengembangkan dan memasarkan *headset*. Selain itu, teknologi yang cepat berubah juga menjadi tantangan dalam mengembangkan dan memasarkan *headset*.

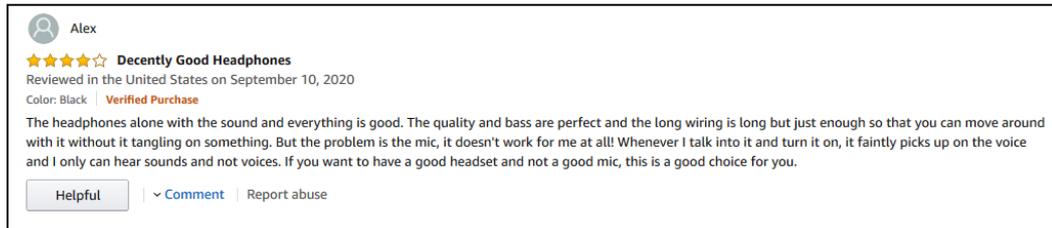
Review produk *headset* yang telah ada dapat dimanfaatkan untuk memahami kebutuhan dan preferensi konsumen yang terus berubah. Karena produk *headset* telah memasuki *e-commerce*, *online reviews* dari produk *headset* dapat digunakan. *Review* produk memberi kesempatan bagi konsumen produk untuk memberi nilai dan komentar terhadap produk yang telah dibeli. Suatu *review* produk umumnya berisi penilaian berupa *rating* dan komentar terhadap suatu

produk. *Rating* yang diberikan konsumen merepresentasikan kepuasan konsumen terhadap produk secara keseluruhan. *Rating* yang tinggi mengindikasikan kepuasan konsumen dan *rating* yang rendah mengindikasikan kurang puasnya konsumen terhadap produk yang dibeli. Dengan demikian, *rating* dapat menunjukkan tingkat kepuasan konsumen. Namun, *rating* belum bisa memberikan informasi secara detail mengenai produk yang dinilai.

Rating merepresentasikan penilaian konsumen terhadap produk tersebut secara umum. Informasi yang bisa didapatkan dari berubahnya *rating* sebatas pada terjadinya perubahan kepuasan konsumen terhadap produk tersebut. Pencarian penyebab perubahan *rating* dapat dilakukan dengan melihat komentar yang menyertai pemberian *rating* terhadap produk tersebut. Dengan menganalisis sentimen pada tiap komentar, dapat diketahui apakah komentar tersebut bersifat positif, negatif, atau netral.

Di dalam suatu komentar, bisa terdapat lebih dari satu opini. Dalam penelitian ini, istilah opini merupakan konsep secara luas yang mencakup sentimen, evaluasi, penilaian, maupun sikap dan informasi yang terkait seperti target opini dan orang yang memberi opini, sedangkan istilah sentimen digunakan dalam arti perasaan positif atau negatif yang mendasari dan tersirat dalam opini (Liu, 2015). Ada kemungkinan terdapat opini bersentimen negatif dan opini bersentimen positif dalam suatu komentar yang sama. Hal ini disebabkan karena di dalam suatu entitas atau produk terdapat aspek-aspek tertentu yang menjadi perhatian konsumen. Oleh karena itu, perlu diidentifikasi aspek-aspek apa saja dari produk yang diperhatikan konsumen dalam menilai produk.

Satu jenis produk bisa memiliki lebih dari satu aspek dan setiap aspek dapat memunculkan opini yang berbeda-beda. Dengan demikian, penilaian terhadap aspek-aspek suatu produk menjadi pertimbangan dalam menilai produk secara keseluruhan dan belum tentu semua aspek mendapat penilaian yang sama. Setiap konsumen memiliki preferensi masing-masing mengenai aspek apa yang lebih diperhatikan dalam melakukan penilaian produk. Untuk mempermudah dalam memahami identifikasi masalah yang sudah disebutkan, Gambar 1.2 menampilkan salah satu *review* dari konsumen sebagai contoh.

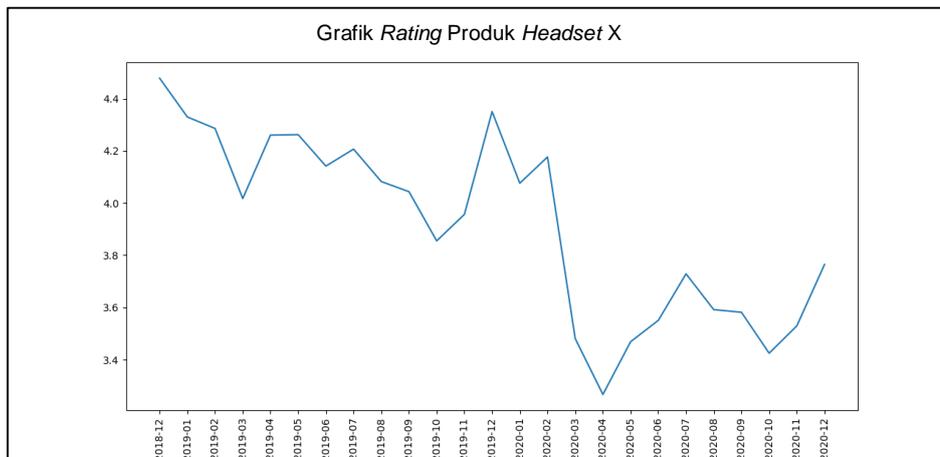


Gambar I.2 Tampilan *Review* di *Website Amazon*

Gambar I.2 merupakan salah satu tangkapan layar dari *review* produk *headset* yang diambil dari *website Amazon*. Dalam *review* tersebut terdapat beberapa komponen seperti nama atau *username* pemberi *review*, *rating* untuk produk secara keseluruhan, judul *review*, tanggal *review*, dan komentar terhadap produk. Pada *review* tersebut, *rating* yang ada merupakan *rating* secara keseluruhan yang menggambarkan penilaian dari konsumen terhadap produk secara keseluruhan. Pada *review* di atas, *rating* yang diberikan adalah 4 dari 5 bintang. Hal ini berarti secara keseluruhan konsumen menilai produk tersebut dengan cukup positif. Namun, jika melihat pada komentar yang diberikan, konsumen memiliki pendapat yang berbeda untuk setiap aspek produk. Untuk aspek suara, kualitas, dan *bass* konsumen memberikan pendapat yang positif. Namun, untuk aspek *mic*, konsumen memberi pendapat yang negatif.

Dari contoh tersebut, dapat diperhatikan bahwa terdapat lebih dari satu aspek pada produk tersebut dan setiap aspek memunculkan opini yang berbeda. Opini penting bagi perusahaan karena opini dapat menunjukkan pandangan konsumen terhadap produk. Sentimen terhadap setiap aspek dapat menunjukkan opini terhadap setiap aspek. Dengan mengetahui sentimen, dapat dilakukan evaluasi terhadap produk untuk melakukan perbaikan pada produk tersebut ataupun mengimplementasikan perbaikan melalui produk baru. Oleh karena itu, perlu diketahui bagaimana sentimen dari konsumen terhadap aspek-aspek yang teridentifikasi.

Berdasarkan pengamatan terhadap *review* untuk salah satu produk *headset* di *website Amazon*, terdapat perubahan penilaian berdasarkan *rating* dari waktu ke waktu. Salah satu contoh perubahan nilai terdapat pada Gambar I.3. Dalam hal ini perubahan yang ada menunjukkan kecenderungan penurunan.



Gambar I.3 Grafik *Rating* Produk *Headset X*

Gambar I.3 adalah grafik *rating headset* untuk produk headset X. Pada Gambar I.3, sumbu x menunjukkan periode dan waktu (Tahun-Bulan), sedangkan sumbu y menunjukkan rata-rata *rating* per bulan yang diberikan oleh konsumen. Dapat dilihat bahwa *rating* cenderung menurun seiring dengan berjalannya waktu sehingga dapat dikatakan bahwa kepuasan konsumen cenderung menurun.

Berdasarkan Gambar I.3, dapat diketahui bahwa *rating* yang diberikan konsumen tidak selalu konstan, melainkan mengalami perubahan seiring waktu. Karena aspek-aspek produk menjadi pertimbangan bagi konsumen dalam memberi penilaian secara keseluruhan, perubahan *rating* seiring waktu dapat mengindikasikan adanya kemungkinan perubahan preferensi konsumen terhadap aspek-aspek produk tersebut seiring berjalannya waktu. Dengan memahami perubahan preferensi konsumen, proses pengembangan produk dapat terbantu untuk menentukan aspek atau atribut apa yang sebaiknya lebih diperhatikan dalam produk yang dikembangkan. Dengan demikian, produk yang dikembangkan dapat lebih sesuai dan memuaskan kebutuhan konsumen. Selain itu, dengan memahami perubahan preferensi konsumen kegiatan pemasaran produk *headset* juga terbantu. Oleh karena itu, selain menentukan aspek-aspek dan melakukan analisis sentimen terhadap aspek-aspek tersebut, juga perlu diketahui apakah preferensi konsumen terhadap aspek-aspek yang teridentifikasi mengalami perubahan seiring berjalannya waktu.

Sebelumnya, terdapat penelitian serupa yang berjudul “Identifikasi *Customer Needs* Terhadap Produk Laptop pada Sebuah Ulasan Produk dengan Metode *Aspect-Based Sentiment Analysis*” (Rachdian, 2020) yang

mengidentifikasi kebutuhan konsumen pada ulasan dengan *aspect based sentiment analysis*. Penelitian tersebut dilakukan terhadap produk *laptop*. Metode *topic modelling* yang digunakan pada penelitian tersebut adalah *Non-negative Matrix Factorization* (NMF). Dalam penelitian ini akan dilakukan *topic modelling* dengan menggunakan metode yang berbeda yaitu *Latent Dirichlet Allocation* (LDA). Metode NMF dan metode LDA merupakan metode yang cukup banyak digunakan untuk *topic modelling*. Ide dasar dari NMF dan LDA adalah sama yaitu bertujuan untuk menemukan topik dari sebaran kata-kata dalam dokumen. Namun, terdapat perbedaan antara LDA dan NMF. LDA merupakan metode probabilistik yang memberi bobot dalam bentuk probabilitas sedangkan NMF merupakan metode non-probabilistik yang memberi bobot dalam bentuk bilangan positif. Pada penelitian ini, hasil metode LDA akan dibandingkan dengan hasil dari metode NMF. Berdasarkan masalah-masalah yang teridentifikasi tersebut, dapat dirumuskan masalah sebagai berikut.

1. Apa saja aspek yang menjadi perhatian dari konsumen berdasarkan *review* produk?
2. Bagaimana sentimen konsumen terhadap aspek-aspek yang teridentifikasi?
3. Apakah terdapat perubahan preferensi konsumen terhadap aspek yang teridentifikasi seiring berjalannya waktu?
4. Bagaimana aspek-aspek yang teridentifikasi dari *review* produk menggunakan LDA bila dibandingkan dengan hasil menggunakan NMF?

I.3 Pembatasan Masalah dan Asumsi Penelitian

Setelah dilakukan identifikasi permasalahan, dilakukan pembuatan batasan dan asumsi. Batasan masalah bertujuan untuk membatasi ruang lingkup penelitian sehingga penelitian yang dilakukan dapat lebih terfokus dan tepat sasaran. Batasan yang digunakan dalam penelitian ini sebagai berikut.

1. *Review* yang digunakan dalam penelitian adalah *online review* berbahasa Inggris dari *website Amazon*.
2. Penelitian tidak memperhitungkan kata-kata negasi dan pertentangan.

Selain batasan masalah di atas, juga digunakan asumsi dalam penelitian ini. Asumsi yang digunakan dalam penelitian ini adalah sebagai berikut.

1. Komponen-komponen dalam *review* dianggap tidak berubah dan relevan sampai penelitian berakhir.
2. Satu kalimat mewakili satu macam aspek atau topik.
3. Kalimat *review* yang tidak relevan (negasi, pertentangan, dan tidak berhubungan) berjumlah sangat sedikit sehingga tidak berpengaruh terhadap hasil.
4. Setiap *review* yang diberikan relevan terhadap produk.

I.4 Tujuan Penelitian

Dalam penelitian ini perlu ditetapkan tujuan dari penelitian untuk menjawab rumusan masalah yang telah ditetapkan. Dengan adanya tujuan penelitian, diharapkan hasil penelitian dapat lebih akurat dalam menjawab permasalahan yang ada. Tujuan dari penelitian yang dilakukan adalah sebagai berikut.

1. Menentukan aspek-aspek yang menjadi perhatian konsumen berdasarkan *review* produk.
2. Menentukan sentimen konsumen untuk aspek-aspek yang diidentifikasi.
3. Menunjukkan perubahan preferensi konsumen terhadap aspek yang teridentifikasi seiring berjalannya waktu.
4. Membandingkan aspek-aspek teridentifikasi dari *review* produk menggunakan LDA bila dibandingkan dengan hasil menggunakan NMF.

I.5 Manfaat Penelitian

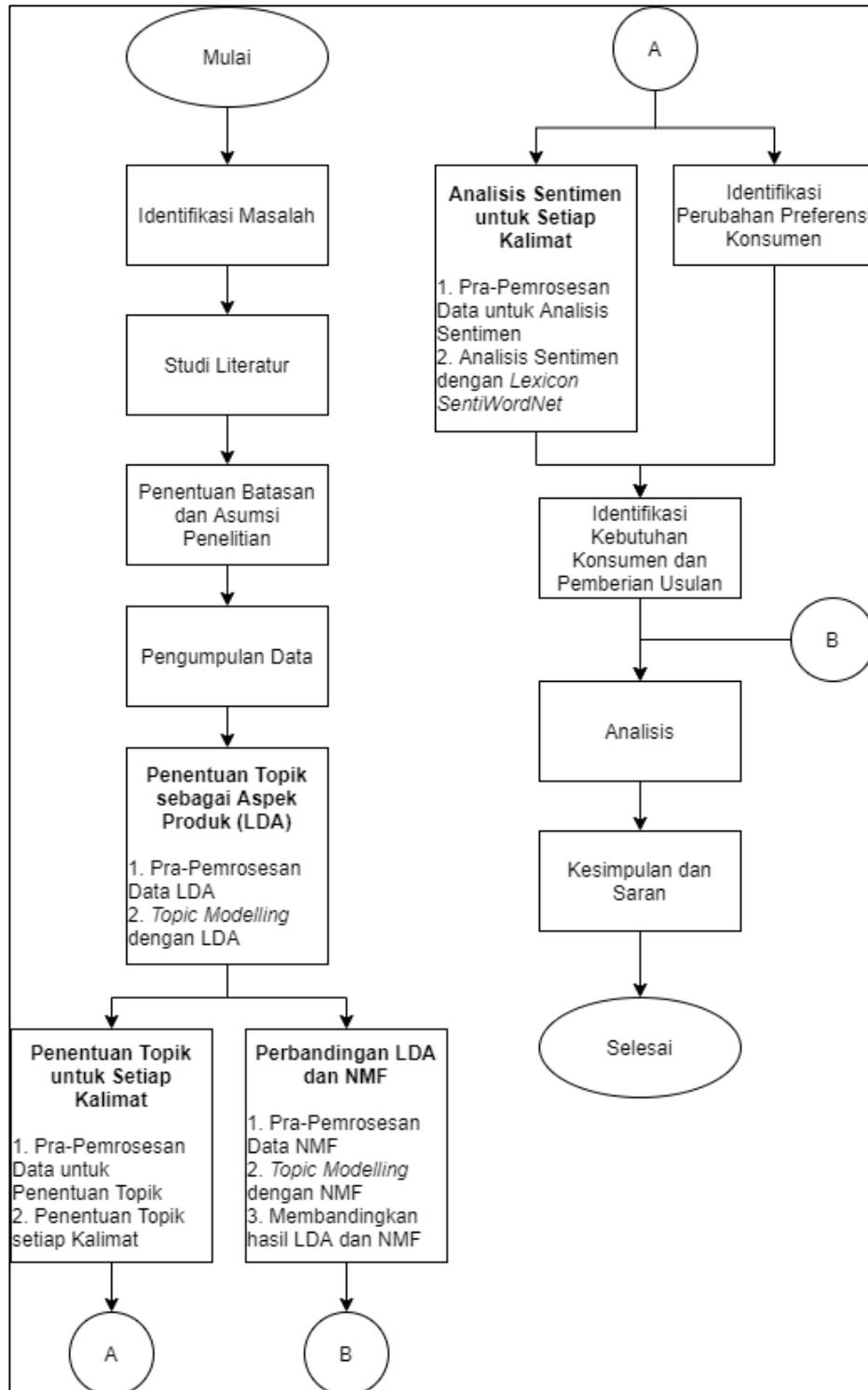
Dalam penelitian yang dilakukan, terdapat manfaat yang dihasilkan bagi pihak yang terkait. Manfaat dari penelitian yang dilakukan adalah sebagai berikut.

1. Hasil penelitian diharapkan dapat memberikan wawasan dan gambaran kepada pembaca mengenai proses pemodelan topik dan analisis sentimen untuk identifikasi kebutuhan dan preferensi konsumen.
2. Hasil penelitian diharapkan dapat digunakan sebagai sarana dan dasar untuk pengembangan produk yang serupa dengan produk yang diteliti.

I.6 Metodologi Penelitian

Metodologi penelitian menggambarkan urutan-urutan hal yang dikerjakan dalam penelitian. Metodologi penelitian ditetapkan agar penelitian yang dilakukan

menjadi lebih sistematis dan lebih terarah. Metodologi penelitian secara keseluruhan dibuat dalam bentuk *flow chart* pada Gambar I.4.



Gambar I.4 *Flow Chart* Metodologi Penelitian

Gambar 1.4 merupakan *flow chart* yang menggambarkan metodologi penelitian secara keseluruhan. Selain *flow chart* metodologi penelitian secara keseluruhan, juga akan dipaparkan secara lebih detail mengenai setiap tahap yang terdapat dalam metodologi penelitian. Berikut merupakan pemaparan tahap-tahap metodologi penelitian secara lebih detail.

1. Identifikasi Masalah

Proses identifikasi masalah bertujuan untuk mendeskripsikan dan merumuskan permasalahan yang ditemukan sebagai langkah awal penelitian. Masalah yang telah teridentifikasi dan dirumuskan mendasari tujuan dari penelitian. Proses ini dilakukan dengan menjelaskan permasalahan yang didapatkan disertai dengan data-data yang menunjukkan masalah tersebut. Dalam hal ini, data yang ditampilkan adalah hasil *plotting* pada *rating* produk terhadap waktu.

2. Studi literatur

Studi literatur dilakukan dengan melakukan studi terhadap teori-teori yang terkait dengan penelitian. Teori-teori tersebut antara lain teori mengenai proses perancangan produk dan pengidentifikasian kebutuhan konsumen, analisis sentimen, *topic modelling*, serta teori-teori lain yang berkaitan dengan metode penelitian.

3. Penentuan batasan dan asumsi penelitian

Dalam penelitian ini juga perlu dilakukan penentuan batasan dan asumsi penelitian. Batasan penelitian bertujuan untuk menciptakan ruang lingkup tertentu pada penelitian sehingga penelitian dapat dilakukan secara lebih terfokus. Asumsi adalah anggapan dasar yang digunakan peneliti dalam konteks penelitian ini. Asumsi penelitian perlu ditentukan karena terdapat kemungkinan adanya pengaruh dari faktor-faktor lain dalam penelitian.

4. Pengumpulan Data

Proses pengumpulan data dilakukan terhadap ulasan *online* pada *website Amazon*. Data yang dikumpulkan meliputi pemberi ulasan, judul ulasan, komentar, *rating*, dan tanggal pemberian ulasan. Data *review* yang diambil adalah data *review* untuk produk *headset* dengan *merk* yang berbeda-beda. Pengumpulan data dilakukan dengan menggunakan bahasa pemrograman *Python* serta *package Selenium*. Data yang dikumpulkan disimpan dalam bentuk tabel dengan format *csv*.

5. Penentuan Topik sebagai Aspek Produk (LDA)

Sebelum melakukan LDA, perlu dilakukan pra-pemrosesan data. Dalam pra-pemrosesan data, dilakukan pembersihan data serta transformasi data. Pembersihan data dilakukan untuk menghilangkan kata atau karakter yang tidak diperlukan. Transformasi data dilakukan untuk mengubah beberapa bentuk data tertentu menjadi bentuk yang lebih sesuai untuk memudahkan pemrosesan data.

Setelah melakukan pra-pemrosesan data, perlu dilakukan penentuan topik yang sesuai berdasarkan kalimat-kalimat *review*. Topik yang didapatkan dapat dianggap sebagai aspek yang diperhatikan konsumen. Aspek yang dimaksud adalah aspek yang dimiliki produk yang bersangkutan dan bukan aspek dari toko atau *seller* seperti pelayanan, pengiriman, dan lain-lain. LDA atau *Latent Dirichlet Allocation* digunakan untuk menentukan topik. Selain metode tersebut, juga akan dilakukan penentuan topik menggunakan *Non-Negative Matrix Factorization* sebagai pembanding. Hasil yang didapatkan dari proses ini adalah aspek-aspek dari produk.

6. Penentuan Topik untuk Setiap Kalimat

Sebelum menentukan topik untuk setiap kalimat, perlu dilakukan pra-pemrosesan data. Dalam pra-pemrosesan data, dilakukan pembersihan data serta transformasi data. Pembersihan data dilakukan untuk menghilangkan kata atau karakter yang tidak diperlukan. Transformasi data dilakukan untuk mengubah beberapa bentuk data tertentu menjadi bentuk yang lebih sesuai untuk memudahkan pemrosesan data seperti pemecahan komentar menjadi kumpulan kalimat.

Topik dari setiap kalimat ditentukan berdasarkan aspek-aspek yang teridentifikasi. Dalam menentukan topik setiap kalimat, digunakan asumsi bahwa setiap kalimat memiliki satu topik. Dengan demikian, setiap kalimat mewakili satu aspek. Penentuan topik dilakukan dengan menentukan topik dengan distribusi terbesar pada kalimat.

7. Analisis Sentimen untuk Setiap Kalimat

Sebelum melakukan analisis sentimen untuk setiap kalimat, perlu dilakukan pra-pemrosesan data. Dalam pra-pemrosesan data, dilakukan pembersihan data serta transformasi data. Pembersihan data dilakukan untuk menghilangkan kata atau karakter yang tidak diperlukan. Transformasi data dilakukan untuk mengubah beberapa bentuk data tertentu menjadi bentuk yang

lebih sesuai untuk memudahkan pemrosesan data seperti pemecahan komentar menjadi kumpulan kalimat.

Setelah aspek telah ditentukan, langkah selanjutnya adalah melakukan analisis sentimen terhadap kalimat-kalimat *review*. Analisis sentimen dilakukan dengan *lexicon based* menggunakan leksikon *SentiWordNet*. Hasil dari analisis sentimen adalah skor sentimen yaitu positif, negatif, dan objektif terhadap setiap aspek. Selain itu, dilakukan rekapitulasi skor sentimen setiap produk.

8. Identifikasi Perubahan Preferensi Konsumen

Pada tahap ini juga akan dilihat apakah terdapat perubahan dari preferensi kebutuhan konsumen di masa lalu dibandingkan dengan kebutuhan konsumen saat ini dengan memperhatikan aspek-aspek produk. Hal ini dilakukan dengan mengidentifikasi perubahan tren dari aspek-aspek yang menjadi perhatian konsumen.

9. Perbandingan LDA dan NMF

Pada tahap ini akan dilakukan perbandingan aspek-aspek yang teridentifikasi dengan LDA dan dengan NMF. Sebelum melakukan perbandingan, dilakukan pra-pemrosesan data untuk NMF. Setelah itu, dilakukan *topic modelling* dengan NMF. Hasil yang didapat menggunakan LDA akan dibandingkan dengan hasil menggunakan NMF.

10. Identifikasi Kebutuhan Konsumen dan Pemberian Usulan

Diakukan identifikasi kebutuhan konsumen melalui interpretasi kalimat setiap aspek. Identifikasi kebutuhan konsumen dilakukan terhadap produk *headset* secara keseluruhan berdasarkan aspek yang didapatkan. Pada tahap ini dilakukan penghasilan usulan. Usulan dihasilkan untuk salah satu produk yang sudah ada dan untuk produk yang sedang dalam pengembangan. Usulan untuk produk yang sudah ada dihasilkan berdasarkan interpretasi kalimat setiap aspek dan sentimen untuk setiap aspek. Usulan untuk produk dalam pengembangan dihasilkan berdasarkan kebutuhan konsumen yang teridentifikasi dan perubahan preferensi konsumen yang teridentifikasi.

11. Analisis

Setelah dilakukan pengolahan data, dilakukan analisis. Analisis dilakukan terhadap proses pengolahan data yang dilakukan sampai tahap pemberian usulan. Selain itu, pada tahap analisis akan dipaparkan hasil penelitian melalui penjelasan. Tahap pengolahan data yang dianalisis meliputi aspek-aspek yang didapatkan,

kebutuhan yang terinterpretasi terhadap aspek, serta perubahan preferensi. Di dalam analisis juga akan dipaparkan mengenai temuan-temuan lain yang berhubungan dengan penelitian.

12. Kesimpulan dan Saran

Pada tahap ini akan diberikan kesimpulan dan saran dari penelitian. Kesimpulan yang didapat dalam penelitian didasarkan pada tujuan penelitian yang ditetapkan sebelumnya. Sedangkan saran yang didapatkan selama penelitian berlangsung dapat menjadi masukan bagi penelitian selanjutnya.

I.7 Sistematika Laporan

Pada subbab ini dijelaskan sistematika penulisan laporan penelitian. Terdapat 5 bab dalam laporan penelitian yang terbagi lagi menjadi subbab-subbab yang lebih detail. Berikut merupakan sistematika penulisan laporan penelitian yang dilakukan.

BAB I PENDAHULUAN

Bab pendahuluan berisi pemaparan latar belakang dan masalah yang mendasari dilakukannya penelitian ini serta tahapan-tahapan penelitian. Oleh karena itu, bab ini terbagi atas latar belakang masalah, identifikasi rumusan masalah, pembatasan masalah, asumsi penelitian, tujuan penelitian, manfaat penelitian, metodologi penelitian, serta sistematika laporan.

BAB II TINJAUAN PUSTAKA

Bab tinjauan pustaka berisi teori-teori yang digunakan dalam penelitian. Teori yang dipaparkan didapat melalui studi pustaka yang sudah ada sebelumnya. Dasar teori yang digunakan meliputi teori mengenai kebutuhan dan preferensi konsumen, *data mining*, *text mining*, metode yang digunakan, serta hal-hal lain yang berkaitan dengan penelitian.

BAB III PENGUMPULAN DAN PENGOLAHAN DATA

Pada bab ini dipaparkan seluruh proses pengumpulan serta pengolahan data. Data yang dikumpulkan merupakan ulasan *online*. Pengolahan data yang dilakukan meliputi pra-pemrosesan data, *topic modelling*, *sentiment analysis*, serta perbandingan metode.

BAB IV ANALISIS

Pada bab ini dilakukan analisis terhadap penelitian yang telah dilakukan. Analisis meliputi analisis terhadap pengumpulan data, pengolahan data, serta hasil

pengolahan data. Pada bab ini temuan-temuan lain dalam penelitian juga dipaparkan.

BAB V KESIMPULAN DAN SARAN

Bab ini berisi kesimpulan dari penelitian yang dilakukan. Kesimpulan dibuat berdasarkan tujuan penelitian yang telah ditetapkan. Selain itu, pada bab ini juga terdapat saran untuk penelitian selanjutnya.