



PENGEMBANGAN METODE EVALUASI LAYANAN
DAN KEPUASAN PENGGUNA DENGAN PENDEKATAN
PEMODELAN TOPIK DAN ANALISIS SENTIMEN

SEMINAR BIDANG KAJIAN

ULFA HIDAYATI
99223113

PROGRAM DOKTOR TEKNOLOGI INFORMASI
UNIVERSITAS GUNADARMA
Juni 2024

Daftar Isi

Daftar Isi	i
1 Pendahuluan	1
1.1 Latar Belakang	1
1.2 Batasan dan Tujuan	3
1.3 Kontribusi	4
2 Tinjauan Pustaka	5
2.1 Kecerdasan Artifisial (<i>Artificial Intelligence</i>) dan Pembelajaran Mesin (<i>Machine Learning</i>)	5
2.2 Pemodelan Topik (<i>Topic Modelling</i>)	6
2.3 Analisis Sentimen	8
2.4 Skala Psikologi	9
2.5 Kepuasan Pengguna	10
2.6 <i>Response Bias</i>	11
2.7 Perbandingan Tinjauan	12
3 Metodologi	16
3.1 Motivasi	16
3.2 Framework Riset	17
3.3 Pendekatan	19
Bibliografi	20

Bab 1

Pendahuluan

1.1 Latar Belakang

Kualitas produk atau layanan dikenal sebagai konstruksi yang memiliki berbagai dimensi dan memiliki tantangan signifikan dalam pengukurannya karena preferensi konsumen yang beragam terhadap berbagai aspek kualitas (Hjorth-Andersen, 1984; Kamakura *et al.*, 1988). Karena karakteristiknya yang tidak berwujud, bervariasi, dan tidak dapat dipisahkan (Parasuraman, Zeithaml, & Berry, 1985), maka layanan memiliki tingkat kompleksitas yang lebih tinggi terutama ketika terjadi kegagalan. Kualitas layanan sangat tergantung pada berbagai aspek dari penawaran layanan tersebut.

Persepsi dan kepuasan pengguna terhadap layanan suatu teknologi sangat berpengaruh pada keberhasilan teknologi baru, sehingga layanan suatu teknologi terbaru hanya akan diadopsi jika pengguna memiliki persepsi positif dan merasa puas dengan penggunaannya (Oliver, 1994; Sun *et al.*, 2009). Kepuasan pengguna teknologi biasanya terbentuk melalui beberapa interaksi dengan sebuah layanan, sedangkan kepuasan penggunaan muncul dalam konteks interaksi tersebut (Wixom dan Todd, 2005).

Kepuasan pengguna biasanya diukur sebagai jumlah atau rata-rata dari kepuasan penggunaan yang berbeda dari berbagai interaksi layanan. Perusahaan biasanya meminta pengguna untuk menilai pengalaman layanan setelah layanan tersebut digunakan, karena hal tersebut membantu dalam meningkatkan layanan dan dapat menghasilkan konten buatan pengguna yang positif atau negatif sebagai pengukuran kepuasan penggunaan (*user satisfaction*). Salah satu cara untuk mendapatkan penilaian pelanggan atau pengguna adalah dengan melakukan survei.

Skala psikologi merupakan salah satu jenis alat ukur yang mengukur atribut non-kognitif (Azwar, 2012). Skala Likert merupakan jenis skala psikologi yang paling banyak digunakan dalam survei dibandingkan skala lainnya, seperti skala Thurstone, skala Guttman, maupun skala diferensial semantik. Sebagian besar skala psikologi

termasuk ke dalam inventori atau *self-report*, di mana responden memberikan respons dikarenakan yang paling mengetahui keadaan sebenarnya dari seseorang adalah dirinya sendiri.

Respons mengenai item atau tingkat nilai dalam survei hanya valid jika tidak ada pengaruh sistematis lainnya terhadap respons item. Namun, seringkali hal ini tidak terjadi dan terdapat faktor tambahan yang mempengaruhi respons item. *Response bias* merupakan respons pada alat ukur yang tidak sesuai dengan keadaan sebenarnya dari responden. Ketika responden diberikan suatu pernyataan survei, maka responden tersebut tidak hanya mencocokkan dengan apa yang ada di dalam dirinya, tetapi juga mempertimbangkan faktor-faktor yang lain seperti ingin dianggap baik, maka faktor-faktor lain inilah yang akan mempengaruhi atribut psikometri dari skala yang digunakan. Validitas skala akan menjadi rendah karena jawaban responden tidak sesuai dengan kenyataannya. Begitu juga dengan reliabilitasnya, jawaban responden menjadi tidak konsisten dalam respons terhadap setiap pernyataan skala. Ketika responden memberikan respons yang tidak sesuai dengan keadaan sebenarnya pada sebuah skala psikologi, maka berdampak kepada keakuratan hasil dari skala itu sendiri. Secara lebih spesifik, *response bias* dapat mempengaruhi variabilitas dari skor tes, reliabilitas, validitas, serta penggunaan dari skor tes tersebut (Shultz, *et al.*, 2014).

Selain itu, kendala yang ditemukan lainnya adalah banyaknya data tidak terstruktur seperti tanggapan terbuka terhadap keluhan pelanggan dan tingkat konsistensi respons pengguna terhadap tanggapan terbuka berupa pertanyaan deskriptif dan tanggapan tertutup berupa pertanyaan pilihan, misalnya memilih “puas” pada pertanyaan tertutup tapi memberikan kritik pada pertanyaan terbuka. Jika tanggapan terbuka itu tidak melalui proses ekstraksi, analisis, dan pengubahan menjadi intervensi yang dapat ditindaklanjuti, maka hal tersebut tidak akan meningkatkan pengambilan keputusan dan kinerja perusahaan (Xu, Frankwick, & Ramirez, 2016). Apabila faktor-faktor yang tidak diketahui membiaskan kepuasan dalam survei, hal tersebut dapat mempengaruhi pengukuran kepuasan secara akurat.

Kemajuan dalam metodologi *big data* dan *text mining* dapat dimanfaatkan untuk mengukur kepuasan pengguna dan mengevaluasi data ulasan atau timbal balik dari survei untuk mendapatkan wawasan berharga, demi mengetahui tingkat keberhasilan serta meningkatkan kualitas baik dari segi pelayanan dan teknologi. Namun, metode yang ada saat ini untuk menentukan indikator evaluasi kualitas seringkali bergantung pada pengambilan keputusan kelompok (Niu, *et al.*, 2021; Rong, *et al.*, 2021) yang mengakibatkan rendahnya kemampuan adaptasi dan penerapan evaluasi menyeluruh.

Berdasarkan latar belakang tersebut, akan dikembangkan model pengukuran kepuasan pengguna dan evaluasi cerdas dan lebih efektif dari sebuah layanan dan teknologi yang diterapkan sebagai penunjang layanan tersebut, dengan memanfaatkan pemodelan topik dan analisis sentimen. Model Pemodelan topik memungkinkan pemodelan untuk mencari ketergantungan antara masing-masing topik dengan

variabel respons dan bagaimana cara penanganan bias respons serta konsistensi respons. Model tersebut dapat mengasumsikan hubungan antar kata. Setelah diidentifikasi, sekelompok kata yang muncul bersamaan akan mewakili topik yang koheren secara tematis. Keluaran dari proses pemodelan topik adalah serangkaian daftar kata yang muncul bersamaan pada kumpulan teks dalam korpus yang sedang dianalisis. Selain itu, akan dianalisis bagaimana perolehan informasi ini dapat menjelaskan lebih lanjut variasi kepuasan pengguna secara keseluruhan dari tema umum yang diungkapkan oleh pengguna. Sehingga data akan dilatih mengenai respons survei pengguna untuk memprediksi kepuasan pengguna.

Dalam penelitian ini, akan dilakukan juga pengukuran kepuasan terhadap sebuah layanan dan sarana teknologi diterapkan pada layanan tersebut. Kepuasan pengguna terhadap layanan diukur melalui proksi sentimen suatu topik yang diidentifikasi melalui ringkasan teks dalam konten buatan pengguna. Topik tersebut mewakili konteks penggunaan dan kata-kata terkait yang relevan dengan konteks tersebut, yang disampaikan oleh pengguna saat menyampaikan pengalaman atau timbal balik dari penggunaan layanan suatu teknologi. Kepuasan pengguna adalah variabel dependen yang diukur menggunakan metode penambangan sentimen dari analisis survei dengan menganalisis polaritas topik.

1.2 Batasan dan Tujuan

1.2.1. Rumusan Masalah

Rumusan masalah berdasarkan latar belakang yang telah dijelaskan sebelumnya adalah:

1. Bagaimana mengembangkan model pengukuran kepuasan pengguna dan evaluasi cerdas dan lebih efektif dari sebuah layanan dan teknologi yang diterapkan sebagai penunjang layanan tersebut, dengan memanfaatkan pemodelan topik dan analisis sentimen?
2. Bagaimana mengidentifikasi *response bias* dan konsistensi respons yang diberikan pengguna dalam survei serta penanganan yang perlu dilakukan?
3. Apa saja fitur kualitas layanan untuk menambah prediktabilitas pada model kepuasan pengguna dan bagaimana fitur-fitur tersebut menangkap aspek kualitas layanan dengan mempertimbangkan dinamika dan saling ketergantungan antar aspek layanan?

1.2.2. Batasan Penelitian

Batasan dari penelitian ini adalah analisis dan pengembangan model pengukuran kepuasan pengguna dan evaluasi cerdas dan lebih efektif dari sebuah layanan dan teknologi yang diterapkan sebagai penunjang layanan tersebut, dengan

memanfaatkan pemodelan topik dan analisis sentimen. Model Pemodelan topik memungkinkan pemodelan untuk mencari ketergantungan antara masing-masing topik dengan variabel respons dan bagaimana cara penanganan bias respons serta konsistensi respons. Selanjutnya, informasi ini akan dianalisa untuk menjelaskan lebih lanjut variasi kepuasan pengguna secara keseluruhan sehingga data akan dilatih mengenai respons survei pengguna untuk memprediksi kepuasan pengguna dan mengevaluasi layanan serta teknologi yang diterapkan.

Subjek penelitian merupakan pengguna dari sebuah layanan yang menerapkan teknologi pada layanan tersebut. Dataset yang diperoleh berasal dari pengumpulan data primer hasil survei dari subjek penelitian.

1.2.3. Tujuan Penelitian

Tujuan dari penelitian ini berdasarkan rumusan masalah yang telah ditetapkan adalah melakukan analisa untuk:

1. Mengembangkan model pengukuran kepuasan pengguna dan evaluasi cerdas dan lebih efektif dari sebuah layanan dan teknologi yang diterapkan sebagai penunjang layanan tersebut, dengan memanfaatkan pemodelan topik dan analisis sentimen.
2. Mengidentifikasi *response bias* dan konsistensi respons yang diberikan pengguna dalam survei serta penanganan yang perlu dilakukan.
3. Mengidentifikasi fitur-fitur kualitas layanan untuk menambah prediktabilitas pada model kepuasan pengguna dan bagaimana fitur-fitur tersebut menangkap aspek kualitas layanan dengan mempertimbangkan dinamika dan saling ketergantungan antar aspek layanan.

1.3 Kontribusi

Kontribusi utama dari penelitian ini adalah untuk membantu dalam mengetahui tingkat keberhasilan suatu layanan dan penerapan teknologi. Selain itu kepuasan pengguna dapat digunakan sebagai alat bantu dalam pengambilan keputusan dan bahan evaluasi untuk perkembangan layanan dan teknologi di masa mendatang. Penelitian ini diharapkan juga dapat menjadi acuan dalam membantu penelitian terkait selanjutnya.

Bab 2

Tinjauan Pustaka

2.1 Kecerdasan Artifisial (*Artificial Intelligence*) dan Pembelajaran Mesin (*Machine Learning*)

Kecerdasan Artifisial atau *Artificial Intelligence* (AI) memiliki sejarah lebih lama dari yang dipahami secara umum dari sains dan filsafat mulai dari Yunani kuno (Dennehy, 2020), tetapi iterasi modernnya sangat bergantung pada Alan Turing (Turing, 1950) dan konferensi di Dartmouth College pada tahun 1956 (McCorduck, 2004), dimana istilah "*Artificial Intelligence*" secara resmi diciptakan dan didefinisikan oleh John McCarthy pada saat itu sebagai "ilmu dan rekayasa pembuatan mesin cerdas".

Salah satu inisial paradigma AI adalah AI berputar di sekitar kognisi tingkat tinggi. Bukan kemampuan untuk mengenali konsep, memahami objek, atau mengeksekusi kompleks keterampilan motorik yang dimiliki, tetapi potensi untuk terlibat penalaran multi-langkah, untuk memahami arti dari bahasa alami, untuk merancang artefak inovatif, untuk menghasilkan rencana baru yang mencapai tujuan, dan bahkan untuk bernalar tentang penalaran mereka sendiri (Langley, 2011).

Untuk AI yang kuat, pendekatan utama berpusat pada simbolik penalaran, bahwa komputer bukan sekadar kalkulator numerik melainkan manipulator simbol umum. Perilaku cerdas muncul membutuhkan kemampuan untuk menafsirkan dan memanipulasi struktur simbolik (Newell dan Simon, 1976).

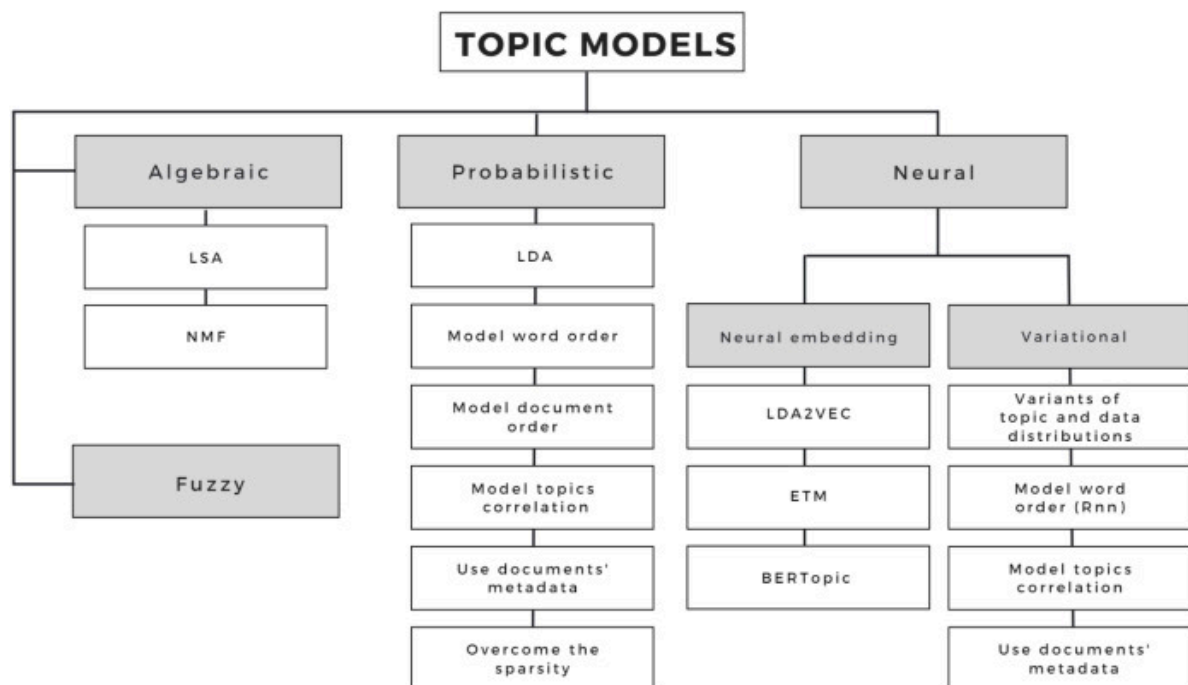
Pembelajaran mesin (*machine learning*) adalah cabang ilmu komputer yang secara luas bertujuan untuk memungkinkan komputer belajar tanpa diprogram secara langsung (Samuel, 1959). *Machine learning* berasal dari gerakan kecerdasan buatan pada tahun 1950-an dan menekankan tujuan dan penerapan praktis, khususnya prediksi dan optimalisasi. Komputer belajar dalam pembelajaran mesin dengan meningkatkan kinerjanya dalam mengerjakan tugas melalui pengalaman. Dalam praktiknya, pengalaman biasanya didapatkan dari menyesuaikan dengan data. Oleh karena itu, tidak ada batasan yang jelas antara pembelajaran mesin dan pendekatan statistik.

Meskipun terdapat kesamaan metodologi, *machine learning* dapat dibedakan

secara filosofis dan praktis. *Machine learning* umumnya menekankan akurasi prediktif dibandingkan inferensi berdasarkan hipotesis, biasanya berfokus pada kumpulan data yang besar dan berdimensi tinggi. Terlepas dari perbedaan yang tepat antara pendekatan-pendekatan tersebut, pembelajaran mesin menawarkan alat-alat penting bagi para ahli epidemiologi. Secara khusus, fokus yang berkembang pada *Big Data* menekankan masalah dan kumpulan data yang mana algoritma *machine learning* lebih unggul daripada pendekatan statistik yang lebih umum digunakan (Rasmussen, 2003; Breiman, 2001).

2.2 Pemodelan Topik (*Topic Modelling*)

Pemodelan topik adalah pendekatan penambangan teks berbasis pembelajaran mesin (*machine learning*), yang bertujuan untuk mengungkap tema dari dokumen teks (Blei, 2012). Meskipun proses pengkodean manusia tetap menjadi standar emas untuk analisis konten tekstual (Short *et al.*, 2010), menganalisis data dalam jumlah besar masih jauh melampaui kapasitas pemrosesan manusia (Kumar dan Zymbler, 2019). Gambar 2.1 menunjukkan algoritma pemodelan topik dan pengaplikasiannya.



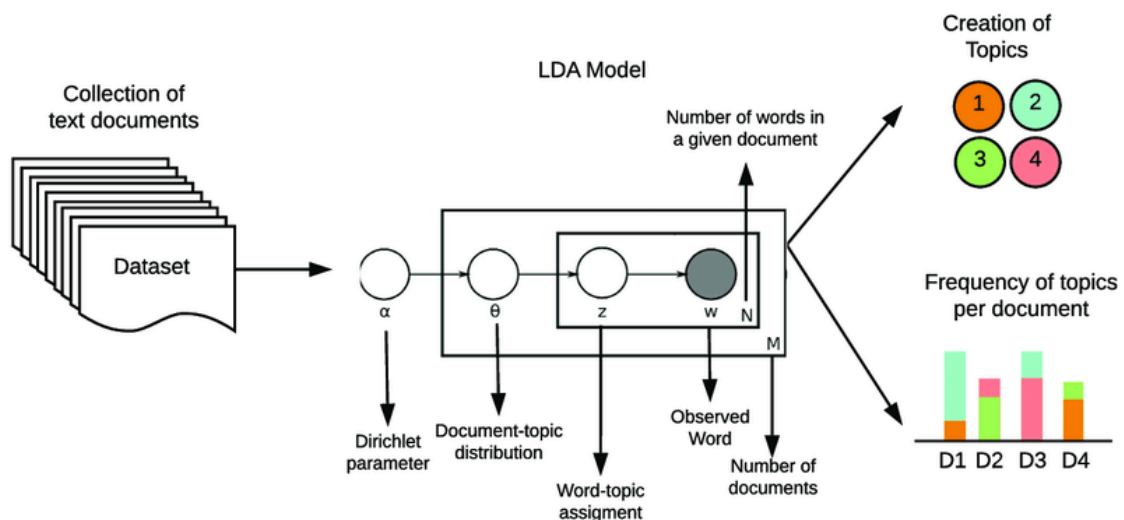
Gambar 2.1. Algoritma Pemodelan Topik dan Pengaplikasiannya

Pemodelan topik memberikan solusi otomatis untuk menganalisis data besar yang tidak terstruktur tersebut tanpa memerlukan pengkodean manual, sehingga mengurangi waktu dan biaya, serta bias manusia. Dibandingkan dengan pendekatan penambangan teks berbasis frekuensi kata yang diterapkan dalam penelitian (Ju *et al.*, 2019; Padma dan Ahn, 2020), pemodelan topik dapat memberikan hasil yang lebih

bermakna, karena pendekatan berbasis frekuensi mengabaikan konteks dan hubungan yang mendasari antara kata-kata (Ahmad dan Larocche, 2015).

Salah satu pemodelan topik yang populer adalah *Latent Dirichlet Allocation* (LDA). LDA melakukan dua tugas yaitu menemukan topik dari korpus dan pada saat yang sama, menetapkan topik-topik tersebut ke dokumen yang ada dalam korpus yang sama. Diagram skema pada Gambar 2.2 merangkum proses kerja LDA dalam menemukan topik.

Penelitian sebelumnya yang dilakukan oleh Sanchez menggunakan LDA untuk menguji hubungan antara berbagai aspek layanan perhotelan dan kualitas layanan (Sanchez-Franco, et al., 2019). Penelitian yang dilakukan Ibrahim dan Wang menggunakan LDA untuk mengekstrak topik utama yang menjadi perhatian pelanggan guna mendukung peningkatan layanan ritel online (Ibrahim dan Wang, 2019). Selain itu, LDA juga memberikan solusi yang baik untuk memahami kepuasan pelanggan di industri berbasis layanan. Lucini menggunakan LDA untuk mengeksplorasi dimensi pelanggan maskapai penerbangan dengan menganalisis 55.000 ulasan pelanggan online (Lucini, et al., 2020).

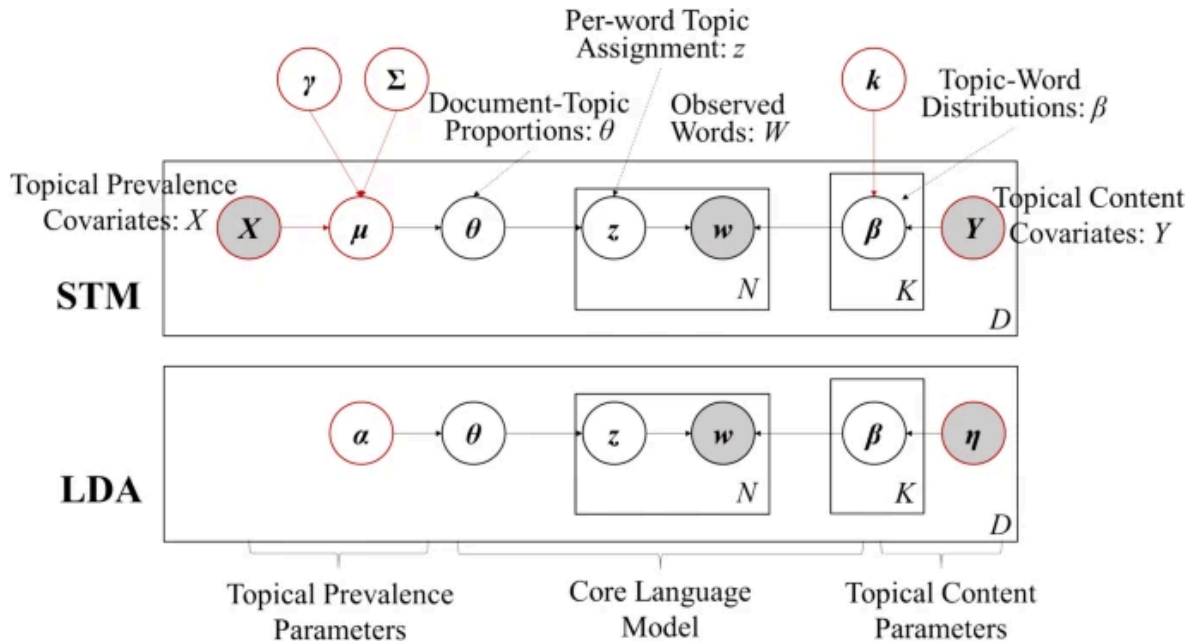


Gambar 2.2. Cara Kerja *Latent Dirichlet Allocation* (LDA)

Selain LDA, ada juga beberapa teknik pemodelan topik lainnya, seperti *Structural Topic Model* (STM) yang dikembangkan oleh Roberts *et al.* (2014), dan *Dynamic Topic Model* (DTM) yang dikembangkan oleh Blei dan Lafferty (2006). Gambar 2.3 menunjukkan perbandingan antara STM dan LDA.

Meskipun semua alat pemodelan topik ini juga dapat mengekstrak topik laten dari teks, alat ini memiliki fokus yang berbeda dan memerlukan penambahan variabel tambahan dalam aplikasi. Misalnya, penekanan penggunaan STM adalah untuk menguji dampak kovariat yang telah ditentukan sebelumnya terhadap perubahan prevalensi topik (Korfiatis et al., 2019). Sedangkan DTM sering diterapkan untuk mengidentifikasi perubahan persepsi pelanggan terhadap produk tertentu dari waktu ke waktu (Ha et

al., 2017). Meskipun tersedia teknik pemodelan topik yang berbeda, pemilihan alat yang tepat harus didasarkan pada model mana yang paling sesuai dengan tujuan penelitian (Saura, 2020).



Gambar 2.3. Perbandingan Antara LDA dan STM

2.3 Analisis Sentimen

Dalam penelitian bisnis, analisis sentimen mengacu pada proses mengidentifikasi emosi yang berbeda (positif, negatif, dan netral) terhadap produk atau layanan dalam teks menggunakan alat pendeteksi sentimen berbantuan komputer (Nasukawa dan Yi, 2003), dan emosi inilah yang sangat berhubungan dengan kepuasan pelanggan. Kumar dan Zymbler (2019) menerapkan pendekatan analisis sentimen berbasis pembelajaran mesin untuk menilai kepuasan pelanggan dari ulasan maskapai penerbangan. Selain itu, Zhu dkk. (2020) meneliti hubungan antara sentimen dan kepuasan tamu di akomodasi Airbnb, mengadopsi metode berbasis leksikon untuk mengklasifikasikan kata-kata dalam setiap ulasan sebagai positif dan negatif, dan mewakili kepuasan tamu menggunakan akumulasi peringkat online. Hasil penelitian ini mengungkapkan bahwa listingan Airbnb cenderung memperoleh skor penilaian lebih tinggi ketika ulasan mereka memiliki tingkat sentimen positif yang lebih tinggi. Geetha dkk. (2017) menguji hubungan antara sentimen ulasan hotel online dan skor penilaian pelanggan yang digunakan untuk mewakili tingkat kepuasan pelanggan dalam penelitian ini. Penelitian ini menyimpulkan bahwa perubahan rating pelanggan dipengaruhi oleh polaritas sentimen review, dan kekuatan pengaruh polaritas sentimen review terhadap variasi rating pelanggan berbeda pada hotel budget dan premium dengan nilai masing-masing sebesar 44 dan 21%.

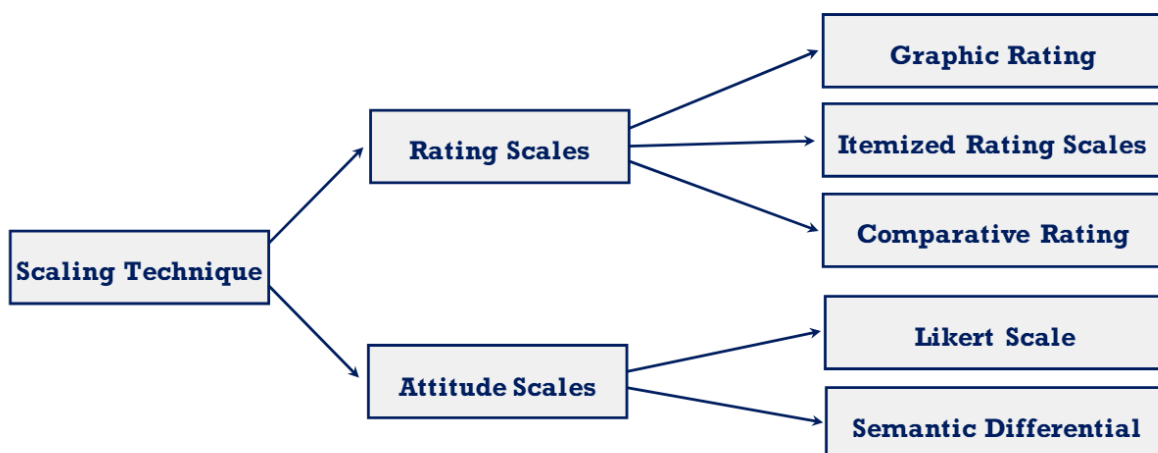
Polaritas sentimen dapat menjadi indikator yang baik untuk kepuasan pelanggan, dan konsep ini juga telah diterapkan dalam penelitian sebelumnya (Xu et al., 2017; Wang et al., 2018).

Dalam hal teknik yang digunakan untuk melakukan analisis sentimen, metode berbasis leksikon berdasarkan daftar fitur leksikal yang telah ditentukan sebelumnya telah diterapkan secara luas dalam penelitian media sosial (Choi et al., 2020). Penerapan metode berbasis leksikon tidak memakan waktu lama dan mudah diterapkan, serta banyak leksikon yang dikembangkan oleh para peneliti di bidang text mining yang dapat langsung diterapkan pada analisis.

2.5 Skala Psikologi

Skala psikologi merupakan salah satu jenis alat ukur yang mengukur atribut non-kognitif (Azwar, 2012). Skala Likert merupakan jenis skala psikologi yang paling banyak digunakan dalam survei dibandingkan skala lainnya, seperti skala Thurstone, skala Guttman, maupun skala diferensial semantik. Sebagian besar skala psikologi termasuk ke dalam inventori atau self-report, di mana responden memberikan respons dikarenakan yang paling mengetahui keadaan sebenarnya dari seseorang adalah dirinya sendiri.

Metode penskalaan dalam survei dibagi menjadi dua kategori utama, yaitu pertanyaan terbuka dan pertanyaan tertutup (Taherdoost, 2017b). Penskalaan adalah proses menghasilkan kontinum, rangkaian yang berkesinambungan nilai-nilai, di mana benda-benda yang diukur ditempatkan. Ada beberapa faktor yang harusnya ada dipertimbangkan untuk memilih metode penskalaan yang sesuai dalam kuesioner. Gambar 2.4 menunjukkan beberapa metode penskalaan yang umum.



Gambar 2.4. Jenis-jenis Teknik Skala Psikologi

Pertanyaan terbuka adalah pertanyaan yang respondennya tidak perlu menyebutkan jawaban tertentu (Taherdoost, 2017a). Pertanyaan terbuka cenderung menghasilkan jawaban yang panjang. Responden kerap kali melihat pertanyaan terbuka sebagai peluang untuk menjawab pertanyaan secara rinci. Berbeda dengan pertanyaan tertutup, responden harus memilih dari sejumlah pertanyaan yang terbatas jawaban potensial (Taherdoost, 2016b). Biasanya jawabannya adalah ya atau tidak. Pertanyaan tertutup mengharuskan responden untuk memilih dari beberapa pilihan jawaban seperti beberapa pertanyaan pilihan, skala Likert dan skala diferensial semantik. Seperti yang diutarakan oleh Taherdoost (2017b), metode skala dapat diklasifikasikan menjadi skala penilaian dan skala sikap.

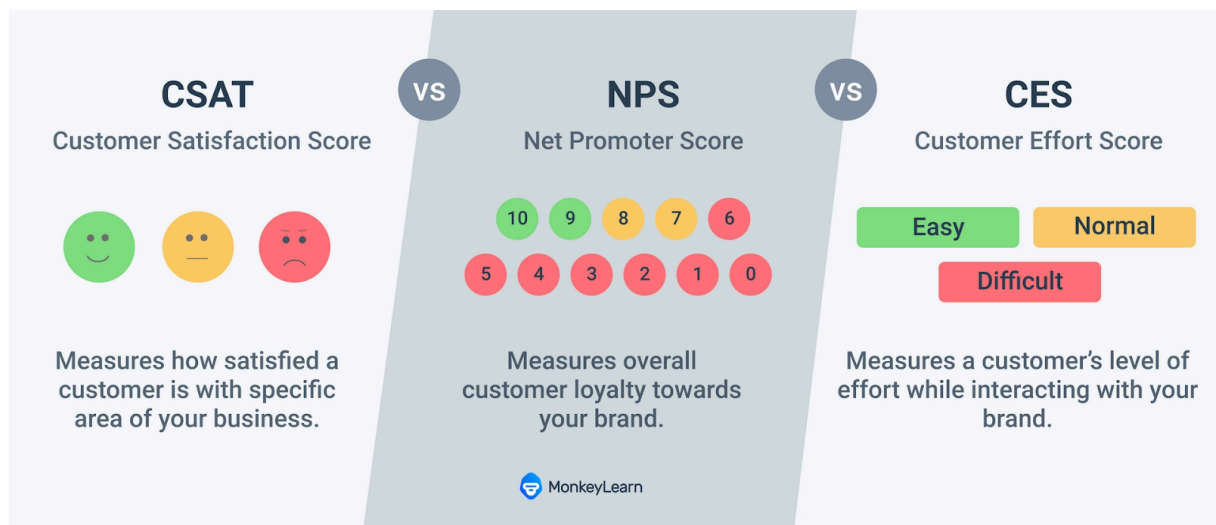
2.6 Kepuasan Pengguna

Survei kepuasan layanan pengguna adalah alat untuk mengumpulkan umpan balik pelanggan mengenai pengalaman mereka dengan sebuah layanan. Survei kepuasan layanan pengguna tentu saja dirancang untuk mengukur tingkat kepuasan pelanggan. Namun, survei ini juga dapat mengidentifikasi area yang perlu ditingkatkan dan mengungkap bagian yang memerlukan perhatian.

Customer Satisfaction Score (CSAT) adalah metrik yang menunjukkan seberapa puas atau tidak puasnya pelanggan atau pengguna terhadap produk dan layanan. Skor CSAT diperoleh melalui survei. Namun, CSAT bukan satu-satunya metrik yang dapat digunakan untuk mengukur kepuasan pengguna. Terdapat berbagai metrik untuk mendapatkan gambaran yang lebih akurat dan menyeluruh tentang tingkat kepuasan pengguna. CSAT paling efektif bila digunakan bersama metrik survei kepuasan pelanggan lainnya seperti *Net Promoter Score* (NPS) dan *Customer Effort Score* (CES). Gambar 2.5 menunjukkan perbedaan di antara ketiganya.

Tak hanya melalui survei, kepuasan pengguna juga dapat diidentifikasi lewat sosial media. Di dunia digital, tujuh puluh persen pelanggan menggunakan situs media sosial untuk mengakses informasi, empat puluh sembilan persen pelanggan menggunakan situs media sosial untuk mengambil keputusan pembelian berdasarkan informasi yang ada di situs media sosial, enam puluh persen pengguna menggunakan media sosial situs hanya untuk berbagi pandangan mereka dengan orang lain; empat puluh lima persen pengguna menggunakan situs media sosial hanya untuk promosi dari mulut ke mulut (Kim dan Ko 2012; Tan dan Lee 2019).

Promosi elektronik dari mulut ke mulut terkait dengan pertemuan layanan dan akses ke internet berdampak pada niat untuk menggunakan layanan yang mendukung teknologi (Chaudrie dan Dwivedi 2005; Ismagilova, et al., 2019). Platform media sosial seperti Facebook dan Twitter memainkan peran penting di dunia digital ini untuk berbagi informasi, menjaga hubungan masyarakat, dan bertukar pendapat antar pengguna (Grover et al., 2017).



Gambar 2.5. Perbedaan *Customer Satisfaction Score* (CSAT), *Net Promoter Score* (NPS), dan *Customer Effort Score* (CES)

Berbagai industri menggunakan situs media sosial untuk memasarkan produk dan layanan mereka seperti industri ritel, pariwisata, berita, hiburan (Grover et al. 2017). Pemasaran dilakukan di situs media sosial untuk produk dan layanan apa pun dengan sedikit usaha dan biaya, juga menjaga hubungan yang kuat dengan pelanggan dan meningkatkan keuntungan industri (Kim dan Ko 2012). Perusahaan juga terlibat dengan pelanggan di media sosial dalam beberapa cara. Pertama adalah dalam bentuk tradisional dimana terdapat komunikasi langsung antara pelanggan dan organisasi melalui promosi berbayar. Bentuk kedua adalah komunikasi sosial di mana pengguna berkomunikasi dengan pengguna lain kemudian mereka mendiskusikan pengalaman pengguna tentang produk dan layanan yang diberikan oleh organisasi dan industri mana pun (Paniagua dan Sapena 2014 ; Mangold dan Faulds, 2009).

2.7 *Response Bias*

Response bias merupakan respons pada alat ukur yang tidak sesuai dengan keadaan sebenarnya dari responden. Secara lebih spesifik, response bias dapat mempengaruhi variabilitas dari skor tes, dan yang pasti reliabilitas, validitas, serta penggunaan dari skor tes tersebut (Shultz, Whitney, & Zickar, 2014).

Berdasarkan penyebabnya, *response bias* dapat dibagi menjadi dua, yaitu *response set* dan *response style*. Beberapa literatur seringkali menganggap keduanya sebagai hal yang sama, namun terdapat perbedaan di antara keduanya. Response set merupakan keinginan sadar maupun tidak sadar sebagai bagian dari seorang responden untuk menjawab dalam cara tertentu untuk menghasilkan gambaran tertentu mengenai dirinya (Rorer, sebagaimana dikutip dalam Ziegler, Griffith, & Gammon, 2015). Response set terjadi setelah responden membaca isi pernyataan lalu membuat respons yang berbeda dari keadaan dirinya. Oleh karena itu, response set

bersifat content-dependent bias (Friedenberg, 1995). Salah satu bentuk response set yang paling dikenal adalah social desirability.

Social desirability terjadi apabila respons seorang responden pada skala dibuat konsisten dengan apa yang ia persepsikan diinginkan oleh orang-orang di sekitarnya (Kuncel & Tellegen, 2009). Bila seorang responden setelah membaca sebuah pernyataan skala, lalu memikirkan bagaimana orang-orang di sekitarnya mengharapkan setiap orang berperilaku terkait perilaku pada skala dan akhirnya memberikan respons pada skala agar sesuai dengan harapan orang-orang tersebut, maka social desirability terjadi. Akibatnya, skor pada skala tidak menggambarkan keadaan sebenarnya dari responden tersebut.

Di lain pihak, response style merupakan kecenderungan seorang responden untuk berespons sistematis terhadap pernyataan skala yang tidak terkait dengan isi pernyataan tersebut (Baumgartner & Steenkamp, sebagaimana dikutip dalam Harzing, 2006). Dibandingkan dengan respons set, response style cenderung lebih konsisten dan stabil dalam berespons (Yang, Harkness, Chin, & Villar, 2010). Dengan kata lain, response style lebih merupakan karakteristik personal dibandingkan karena faktor situasional. Beberapa response bias yang termasuk response style adalah extreme response styles (ERS), kecenderungan responden untuk memilih respons yang ekstrim pada skala (misal, “sangat setuju” atau “sangat tidak setuju”), dan middle response styles (MRS), kecenderungan responden untuk memilih respons di tengah pada skala (misal, “netral” atau “ragu-ragu”) (Harzing, 2006). Sama seperti pada response set, skala Likert juga lebih rentan terjadinya response style (Yulianto, 2019).

Dikarenakan response set terjadi terkait isi dari pernyataan skala, maka agar tidak terjadi response set, salah satu yang dapat dilakukan adalah dengan membuat pernyataan pernyataan sebaik mungkin. Misalnya dengan mengikuti pedoman penulisan pernyataan dari Edwards (sebagaimana dikutip dalam Azwar, 1995) “jangan menulis pernyataan yg akan disetujui/tidak setuju oleh sebagian besar responden” ditujukan untuk menghindari terjadinya social desirability. Berbeda dengan response set, response style tidak dapat dihindari dengan cara membuat pernyataan yang baik, mengingat response bias ini merupakan karakteristik individual yang tidak terkait dengan isi pernyataan skala. Response style lebih mudah dikenali setelah skala terisi.

Dengan demikian, untuk menjaga keakuratan hasil pengukuran, maka data dari skala yang terindikasi memiliki response style sebaiknya tidak disertakan dalam analisis. Misalnya, seorang responden memilih respons “netral” pada lebih dari 75% pernyataan yang diberikan, maka ini dapat menjadi indikasi adanya response style pada responden tersebut sehingga partisipan ini dapat diabaikan responsnya. Namun hal ini harus dilakukan secara hati-hati karena apabila cukup banyak responden yang terindikasi memiliki response style, bisa jadi dikarenakan proses penyusunan skala yang kurang baik.

2.8 Perbandingan Tinjauan

Proses membandingkan dari masing-masing tinjauan, dan pilihan mana yang digunakan atau diadaptasi dapat dilihat pada Tabel 2.1.

Tabel 2.1. Daftar Tinjauan Penelitian Terkait.

JUDUL	NAMA PENULIS	IDENTITAS	MASALAH	SOLUSI	HASIL
What Affects Usage Satisfaction in Mobile Payments? Modelling User Generated Content to Develop the “Digital Service Usage Satisfaction Model”	Arpan Kumar Kar	Inf Syst Front 23, 1341–1361 (2021). https://doi.org/10.1007/s10796-020-10045-0	Layanan pembayaran seluler menjadi semakin penting dalam kehidupan sehari-hari di India karena berbagai peristiwa yang direncanakan dan tidak direncanakan.	Identifikasi faktor-faktor penentu kepuasan penggunaan pembayaran seluler yang dapat meningkatkan adopsi layanan. “Model Kepuasan Penggunaan Layanan Digital” telah diusulkan dan divalidasi dengan menggabungkan adopsi teknologi dan literatur ilmu layanan.	Model penelitian inferensial diusulkan, yaitu DSUSM. Tahap pengembangan model berdasarkan topik yang diidentifikasi dari ringkasan teks tweet. Dengan analisis inferensial menggunakan regresi, faktor-faktor seperti biaya, kegunaan, kepercayaan, kredibilitas, pengaruh sosial, privasi informasi dan daya tanggap telah diidentifikasi memiliki pengaruh yang signifikan terhadap penerapan layanan pembayaran seluler.
Measuring Service Quality from Unstructured Data: A Topic Modeling Application on Airline Passengers’ Online Reviews	Nikolaos Korfiatis, Panagiotis Stamolampros, Panos Kourouthanassis, Vasileios Sagiadinis	Expert Systems with Applications, Volume 116, 2019, Pages 472-486, ISSN 0957-4174, https://doi.org/10.1016/j.eswa.2018.09.037 .	Kualitas layanan adalah sebuah multi-konstruksi dimensi yang tidak diukur secara akurat oleh aspek-aspek yang berasal dari peringkat numerik dan bobot terkait. Kumpulan data besar diperoleh melalui ulasan online penumpang	Memanfaatkan ulasan online yang diidentifikasi dari pemodelan topik data tidak terstruktur untuk memahami kepuasan pelanggan dan selanjutnya kualitas layanan. Penelitian ini menggunakan Structural Topic Models (STM), yang merupakan	Inovasi ini memungkinkan untuk menangkap pendorong kepuasan yang dominan beserta dinamika dan saling ketergantungannya. Hasilnya mengungkap ortogonalitas aspek persaingan maskapai penerbangan berbiaya rendah ketika semua dimensi kualitas layanan lainnya

JUDUL	NAMA PENULIS	IDENTITAS	MASALAH	SOLUSI	HASIL
			maskapai penerbangan dari TripAdvisor	perluasan probabilistik terkini dari Latent Dirichlet Allocation (LDA).	dipertimbangkan, sehingga menjelaskan keberhasilan maskapai penerbangan berbiaya rendah di pasar penerbangan.
Leveraging Text Mining and Analytic Hierarchy Process For The Automatic Evaluation of Online Courses.	Xieling Chen, Haoran Xie, Xiaohui Tao, Fu Lee Wang & Jie Cao	Int. J. Mach. Learn. & Cyber. (2024). https://doi.org/10.1007/s13042-024-02203-6	Metode evaluasi kualitas kursus online dan bobotnya sering kali bergantung pada pengambilan keputusan kelompok. Evaluasi simulasi ahli dan penggunaan indikator yang telah ditentukan sebelumnya mungkin tidak secara akurat mencerminkan pengalaman dan kebutuhan peserta didik, sehingga mengakibatkan kurangnya keterpusatan pada peserta didik dalam evaluasi kursus.	Memperkenalkan metodologi pengambilan keputusan multi-kriteria yang memanfaatkan penambangan teks dan proses hierarki analitik (AHP)	Mengembangkan pendekatan baru untuk mengevaluasi <i>Massive Online Open Courses</i> (MOOC) secara otomatis menggunakan model pengambilan keputusan multi-kriteria, menggabungkan penambangan teks dan AHP.
Exploring Sources of Satisfaction and Dissatisfaction in Airbnb Accommodation Using Unsupervised and Supervised Topic Modeling	Kai Ding, Wei Chong Choo, Keng Yap Ng, Siew Imm Ng, Pu Song	Front. Psychol., 21 April 2021 Sec. Organizational Psychology Volume 12 - 2021 https://doi.org/10.3389/fpsyg.2021.659481	Industri pariwisata dan perhotelan telah mengalami perubahan yang signifikan. Salah satu perubahan signifikannya adalah semakin banyak pelanggan yang secara aktif memperoleh dan berbagi informasi	Periksa atribut-atribut utama yang memengaruhi kepuasan dan ketidakpuasan pengguna Airbnb melalui analisis ulasan online menggunakan pendekatan <i>Latent Dirichlet Allocation</i> (LDA)	Hasil analisis distribusi topik menunjukkan bahwa di berbagai jenis properti Airbnb, pengguna Airbnb memberikan kepentingan berbeda pada atribut layanan yang sama. Analisis korelasi topik mengungkapkan bahwa pengalaman

JUDUL	NAMA PENULIS	IDENTITAS	MASALAH	SOLUSI	HASIL
			mengenai produk dan layanan di platform sosial online dibandingkan hanya mengandalkan informasi yang disediakan oleh penyedia layanan.	dan <i>Supervised LDA</i> (sLDA).	seperti di rumah dan bantuan dari tuan rumah berhubungan dengan niat berkunjung kembali pengguna Airbnb.
What Is the Best Response Scale for Survey and Questionnaire Design; Review of Different Lengths of Rating Scale / Attitude Scale / Likert Scale	Hamed Taherdoost	International Journal of Academic Research in Management (IJARM) Vol. 8, No. 1, 2019, Page: 1-10, ISSN: 2296-1747	Hasil analisis distribusi topik menunjukkan bahwa di berbagai jenis properti Airbnb, pengguna Airbnb memberikan kepentingan yang berbeda pada atribut layanan yang sama. Analisis korelasi topik mengungkapkan bahwa pengalaman seperti di rumah dan bantuan dari tuan rumah berhubungan dengan niat mengunjungi kembali pengguna Airbnb.	Memberikan gambaran umum tentang skala Likert dan membandingkan skala penilaian dengan panjang yang berbeda.	Keputusan mengenai jumlah poin skala Likert yang digunakan untuk survei dan kuesioner mereka. Secara keseluruhan penelitian ini menyarankan penggunaan tujuh poin skala penilaian dan jika ada kebutuhan untuk mengarahkan responden pada satu sisi, maka skala enam poin mungkin bisa digunakan menjadi yang paling cocok.

Bab 3

Metodologi

3.1 Motivasi

Penelitian mengenai pengukuran kepuasan dan khususnya mengenai bagaimana berbagai aspek survei dapat membiaskan kepuasan pengguna masih relatif sedikit. Penelitian pengukuran kepuasan pengguna belum banyak yang memperhatikan bias dan konsistensi antara pertanyaan tertutup dan terbuka dari ungkapan yang diberikan oleh pengguna. Kebanyakan dari pemodelan topik mengambil sumber dari ulasan *review* aplikasi digital dan ulasan pada media sosial. Pemahaman mendalam mengenai karakteristik psikologis manusia dalam memberikan respons dapat diterapkan pada model yang akan dikembangkan dan kemampuan model dalam melatih data akan lebih baik sehingga meningkatkan akurasi prediksi serta keputusan. Selain itu, metode evaluasi layanan yang ada belum memberikan kemampuan adaptasi dan penerapan evaluasi menyeluruh.

Usulan model pengukuran kepuasan pengguna dan evaluasi digunakan untuk mengembangkan sistem cerdas dan lebih efektif dari sebuah layanan dan teknologi yang diterapkan sebagai penunjang layanan tersebut. Metode pemodelan topik digunakan untuk mengontrol jumlah kata dan topik yang muncul dari timbal balik yang diberikan pada survei, sehingga memungkinkan pengembangan model empiris. Selanjutnya akan dilakukan evaluasi untuk melihat bagaimana topik-topik tersebut saling berhubungan. Analisis sentimen dilakukan untuk mengidentifikasi polaritas topik dengan pendekatan semantik. Data akan dilatih mengenai respons survei pengguna untuk memprediksi kepuasan pengguna. Pengukuran kualitas layanan akan dilakukan menggunakan data terstruktur dan tidak terstruktur, kemudian mengubahnya menjadi wawasan manajerial. Sehingga membuka jalan bagi pengembangan sistem cerdas yang lebih efektif.

Fitur kualitas layanan yang diambil dari data tidak terstruktur yang berpotensi menambah prediktabilitas pada model kepuasan pengguna juga akan dianalisis untuk menjawab pertanyaan apakah fitur-fitur tersebut telah mewakili aspek kualitas kepuasan pengguna dengan mempertimbangkan dinamika dan saling ketergantungan antar aspek layanan. Kemudian seluruh informasi yang diperoleh digunakan untuk

mengevaluasi layanan dan teknologi yang diterapkan.

Model proses untuk seluruh analisis yang dilakukan dalam penelitian ini menggunakan model *Cross-Industry Standard Process for Data Mining* (CRISP-DM). Tahapan yang akan dilakukan yaitu *Business Understanding*, *Data Understanding*, *Data Preparation*, *Modeling*, *Evaluation*, dan *Deployment*.

3.2 **Framework Riset**

Penelitian menggunakan metode studi literatur untuk: 1) Memperoleh informasi melalui referensi dari jurnal-jurnal penelitian yang terkait; 2) Mencari definisi dan teori untuk keperluan tinjauan pustaka serta; 3) Mencari informasi mengenai teori survei, skala psikologi (psikometri), response bias, pemodelan topik, analisis sentimen, dan fitur kualitas layanan yang diambil dari data tidak terstruktur yang berpotensi menambah prediktabilitas pada model kepuasan pengguna untuk mengetahui fitur-fitur yang mewakili aspek kualitas kepuasan pengguna dengan mempertimbangkan dinamika dan saling ketergantungan antar aspek layanan.

Berdasarkan tahapan yang ditunjukkan pada Gambar 3.1, tahapan penelitian dapat diuraikan sebagai berikut:

1. Studi Literatur

Tahap pencarian data dan informasi secara mendetail terkait dengan teori survei, skala psikologi (psikometri), response bias, pemodelan topik, analisis sentimen, dan fitur kualitas layanan yang diambil dari data tidak terstruktur yang berpotensi menambah prediktabilitas pada model kepuasan pengguna untuk mengetahui fitur-fitur yang mewakili aspek kualitas kepuasan pengguna dengan mempertimbangkan dinamika dan saling ketergantungan antar aspek layanan. Pencarian informasi mengenai pemodelan topik dan analisis semantik dilakukan untuk membuat sebuah pemodelan kepuasan pengguna agar sasaran dan tujuan dapat tercapai.

2. Pengumpulan Data

Dari informasi yang diperoleh, penelitian berlanjut pada pengumpulan data awal yaitu berupa data primer yang diperoleh melalui survei pada subjek penelitian, mempelajari fitur-fitur apa saja yang akan digunakan, serta mengidentifikasi masalah yang berkaitan dengan kualitas data.

3. Penentuan Algoritma atau Metode Penelitian

Penelitian berlanjut pada tahap penentuan algoritma atau metode pemodelan topik dan analisis semantik dilakukan untuk membuat sebuah pemodelan kepuasan pengguna untuk penelitian, menentukan parameter dengan nilai yang optimal, serta menentukan *tools* yang akan digunakan.

4. Proses Pelatihan Data

Proses pelatihan dilakukan untuk mengembangkan algoritma atau metode

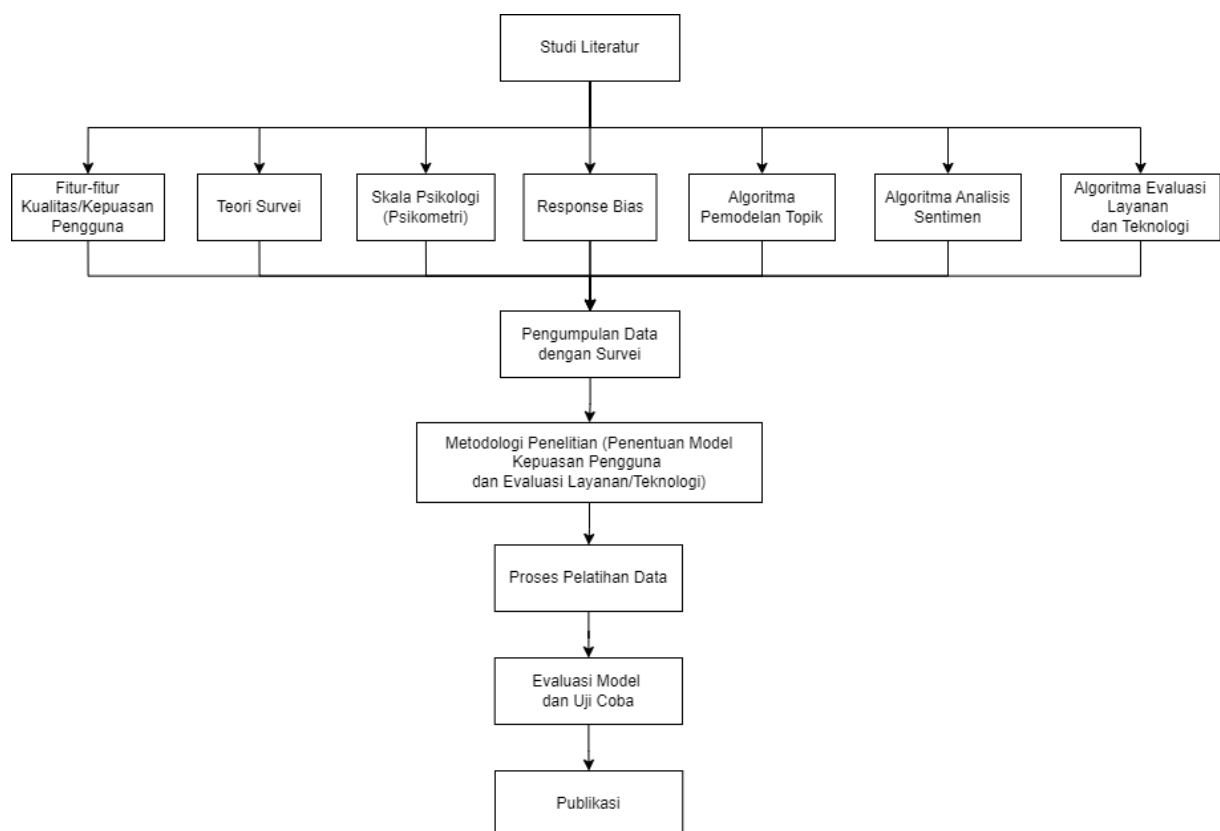
pemodelan topik dan analisis semantik yang membangun model berdasarkan data-data pelatihan untuk membuat prediksi atau pemodelan kepuasan pengguna tanpa diprogram secara eksplisit untuk melakukannya.

5. Evaluasi Model dan Uji Coba

Evaluasi dilakukan untuk mengukur akurasi dari model yang dikembangkan. Hasil dari galat yang diperoleh dari beberapa metode uji.

6. Publikasi

Langkah terakhir adalah fase penyusunan laporan atau presentasi dari hasil dan pembahasan yang didapat dalam penelitian.



Gambar 3.1. Tahapan Penelitian

Kerangka kerja atau struktur konseptual yang digunakan untuk mengarahkan dan mendukung proses penelitian dapat dilihat pada Gambar 3.2.



Gambar 3.2. Alur Metodologi Penelitian

Dataset yang terdiri dari fitur-fitur aspek kepuasan pengguna akan dianalisis menggunakan metode pemodelan topik. Untuk mengatasi *response bias* pada dataset

hasil survei akan dilakukan perluasan analisa pada tahap *Data Understanding* dan *Data Preparation*. Kemudian data hasil pemrosesan akan memasuki tahap pemodelan topik. Keluaran dari pemodelan topik berupa kata-kata yang saling terkait. Dari hasil yang diperoleh akan dilakukan analisis semantik untuk mengecek polaritas topik. Data yang dilatih akan menangani respons survei pengguna untuk memprediksi kepuasan pengguna dan kumpulan topik akan dianalisis untuk mengevaluasi layanan dan teknologi yang diterapkan.

3.3 Pendekatan

Pendekatan metodologi dalam penelitian ini terdiri dari beberapa jenis, di antaranya:

1. Pendekatan Kuantitatif dan Kualitatif

Pendekatan kuantitatif berfokus pada pengumpulan dan analisis data untuk mengukur fitur data dari hasil survei dan mencari hubungan atau perbedaan yang dapat digeneralisasikan dengan menggunakan metode statistik. Proses ini juga dilakukan dengan pengukuran skala psikologi (psikometri). Sedangkan pendekatan kualitatif berfokus pada pengumpulan dan analisis data non-numerik seperti umpan balik pada survei pertanyaan terbuka deskriptif berupa teks. Tujuannya adalah untuk memahami fenomena dari perspektif subjek penelitian.

2. Pendekatan Eksperimental untuk Evaluasi dan Uji Coba

Pendekatan ini melibatkan manipulasi satu atau lebih variabel independen dan mengamati efeknya pada variabel dependen. Fitur-fitur kualitas layanan yang mewakili aspek kualitas kepuasan pengguna merupakan variabel independen sedangkan hasil kepuasan pengguna merupakan variabel dependen.

Penelitian akan menggunakan pendekatan pemodelan topik dan pendekatan aspek psikologi dalam penanganan response bias. Model akan mengasumsikan hubungan antar kata apa pun yang sering muncul bersamaan dalam teks yang sama. Setelah diidentifikasi, sekelompok kata yang muncul bersamaan kemudian diambil untuk mewakili topik yang koheren secara tematis. Keluaran prinsip dari proses pemodelan topik adalah serangkaian daftar kata yang muncul bersamaan dalam kumpulan teks. Setiap daftar kata mewakili topik tersendiri dan kata-kata yang paling khas dari topik tersebut. Kata-kata yang ditampilkan dalam daftar kata suatu topik dapat dipandang sebagai kata kunci dari sebuah teks. Jumlah topik yang disediakan oleh model akan mempengaruhi sifat topik yang dihasilkan. Sejumlah besar topik memberikan tema yang lebih terperinci, sementara sejumlah kecil topik memberikan tema yang lebih umum atau luas.

Metodologi penelitian menggabungkan analisis teks menggunakan penambangan sentimen dan pemodelan topik. Setelah pemodelan telah dilakukan, akan dilakukan proses pencarian konteks teks dan memetakan skor ke tema relevan. Sebuah

kata dapat menjadi satuan ukuran yang dapat dipetakan ke suatu tema. Kemudian perlu dilakukan evaluasi bagaimana topik-topik tersebut saling berhubungan satu sama lain. Informasi yang diperoleh juga digunakan untuk mengevaluasi layanan dan teknologi yang diterapkan. Analisis lebih lanjut digunakan untuk mengetahui jumlah cluster dan penentuan cluster untuk kepuasan pengguna. Tahap selanjutnya adalah mengevaluasi sentimen topik-topik tersebut menggunakan analisis polaritas. Data yang dilatih akan menangani respons survei pengguna untuk memprediksi kepuasan pengguna. Untuk mengatasi response bias pada dataset hasil survei akan dilakukan perluasan arsitektur jaringan saraf sederhana yang memprediksi kepuasan pengguna.

Bibliografi

- Brookes, G., & McEnery, T. (2019). The utility of topic modelling for discourse studies: A critical evaluation. *Discourse Studies*, 21(1), 3-21.
- Ding, K., Choo, W. C., Ng, K. Y., Ng, S. I., & Song, P. (2021). Exploring sources of satisfaction and dissatisfaction in Airbnb accommodation using unsupervised and supervised topic modeling. *Frontiers in psychology*, 12, 659481.
- Janiesch, Christian et al. (2021). *Machine Learning and Deep Learning*. Springer. *Electronic Markets* (2021) 31:685–695. Published online: 8 April 2021.
- Kar, A.K. (2021). What Affects Usage Satisfaction in Mobile Payments? Modelling User Generated Content to Develop the “Digital Service Usage Satisfaction Model”. *Inf Syst Front* 23, 1341–1361. <https://doi.org/10.1007/s10796-020-10045-0>.
- Korfiatis, N., Stamolampros, P., Kourouthanassis, P., & Sagiadinos, V. (2019). Measuring service quality from unstructured data: A topic modeling application on airline passengers’ online reviews. *Expert Systems with Applications*, 116, 472-486.
- Nie Y, Luo H, Sun D (2021) Design and validation of a diagnostic MOOC evaluation method combining AHP and text mining algorithms. *Interact Learn Environ* 29:315–328. <https://doi.org/10.1080/10494820.2020.1802298>.
- Rong L, Wang L, Liu P, Zhu B (2021) Evaluation of MOOCs based on multigranular unbalanced hesitant fuzzy linguistic MABAC method. *Int J Intell Syst* 36:5670–5713. <https://doi.org/10.1002/int.22526>.
- Taherdoost, H. (2019). What is the best response scale for survey and questionnaire design; review of different lengths of rating scale/attitude scale/Likert scale. *Hamed Taherdoost*, 1-10.
- Yulianto, A. (2020). Mewaspadaai response bias dalam skala psikologi. *Buletin KPIN*, 6(3).
- Oliver, R. L. (1994). Conceptual issues in the structural analysis of consumption emotion, satisfaction, and quality: Evidence in a service setting. *Advances in Consumer Research*, 21, 16–22.
- Shultz, K. S., Whitney, D. J., & Zickar, M. J. (2014). *Measurement theory in action : case studies and exercises*. New York: Taylor & Francis.
- Xu, Z., Frankwick, G. L., & Ramirez, E. (2016). Effects of big data analytics and

- traditional marketing analytics on new product success: A knowledge fusion perspective. *Journal of Business Research*, 69 (5), 1562–1566.
<https://doi.org/10.1016/j.jbusres.2015.10.017>
- Lin, G. T., & Sun, C. C. (2009). Factors influencing satisfaction and loyalty in online shopping: an integrated model. *Online Information Review*, 33(3), 458–475.
- Wixom, B. H., & Todd, P. A. (2005). A theoretical integration of user satisfaction and technology acceptance. *Information Systems Research*, 16(1), 85–102.
- Azwar, S. (1995). *Sikap manusia: Teori dan pengukurannya*. Yogyakarta: Pustaka Belajar.
- Azwar, S. (2012). *Penyusunan skala psikologi*. Yogyakarta: Pustaka Pelajar.
- Friedenberg, L. (1995). *Psychological testing: Design, analysis, and use*. Boston: Allyn & Bacon.
- Harzing, A.-W. (2006). Response styles in cross-national survey research: A 26-country study. *International Journal of Cross Cultural Management*, 6(2), 243–266. doi:10.1177/1470595806066332
- Kuncel, N. R., & Tellegen, A. (2009). A conceptual and empirical reexamination of the measurement of the social desirability of items: Implications for detecting desirable response style and scale development. *Personnel Psychology*, 62(2), 201–228. the likelihood.
- Moors, G., Kierulff, N. D., & Vermunt, J. K. (2014). The effect of labeling and numbering of response scales on response bias. *Sociological Methodology*, 44(1), 369–399. doi:<https://doi.org/10.1177/0081175013516114>.
- Yang, Y., Harkness, J. A., Chin, T.-Y., & Villar, A. (2010). Response styles and culture. In Harkness, Survey methods in multinational, multiregional, and multicultural contexts (pp. 203–223). New York: John Wiley & Sons, Inc.
- Yulianto, A. (2019). *Penyusunan skala Guttman untuk pengukuran psikologi*. Tangerang Selatan: Universitas Pembangunan Jaya.
- Ziegler, M., Maaß, U., Griffith, R., & Gammon, A. (2015). What Is the nature of faking? Modeling distinct response patterns and quantitative differences in faking at the same time. *Organizational Research Methods*, 18(4), 679–703. doi:10.1177/1094428115574518
- Samuel, A. L. (1959). Some studies in machine learning using the game of checkers. *IBM Journal of research and development*, 3(3), 210–229.
- Rasmussen, C. E. (2003). Gaussian processes in machine learning. In Summer school on machine learning (pp. 63–71). Berlin, Heidelberg: Springer Berlin Heidelberg.