

BAB II TINJAUAN PUSTAKA

Pandemi Covid-19 yang terjadi menjadi tantangan bagi pemilik bisnis. Carracedo pada tahun 2020 melakukan analisis teks pada *paper* atau jurnal yang membahas tentang pengaruh pandemic terhadap bisnis [12]. Penelitian ini menggunakan metode *Clustering Chi-Squared* untuk menganalisis kumpulan teks tersebut. Lalu dari proses analisis tersebut, Ditemukan kata-kata yang sering muncul dalam *paper* yang dianalisis dan dapat divisualisasikan menjadi tiga *cluster* topik utama, yaitu pengaruh pandemic terhadap manajemen kesehatan, perubahan dalam kebiasaan konsumsi masyarakat, dan pengaruh pandemic terhadap organisasi bisnis. Analisis teks juga diterapkan untuk melakukan analisis sentiment terhadap *review* dari pelanggan terhadap sebuah restoran.

Analisis sentiment pada *review* konsumen dilakukan oleh Gan pada tahun 2016 menggunakan Yelp Dataset untuk dianalisis menggunakan analisis sentiment multidimensional. Hasil analisis menunjukkan bahwa konsumen merasa positif tentang makanan, pelayanan, suasana, dan harga dari restoran yang dikunjungi [13]. Rosetti pada tahun 2014 juga melakukan analisis pada ulasan pengguna yang didapatkan dari basis YELP *dataset* dan TripAdvisor. Makalah ini mengeksplorasi penerapan metode model topik dalam domain pariwisata. Makalah ini menggunakan metode *Topic Model* dan *Topic-Criteria Model*. Skenario aplikasi yang diterapkan dalam pendekatan berupa prediksi *rating* dan rekomendasi, analitik dan interpretasi, serta usulan *rating* dari sebuah ulasan [14].

Penerapan lain dari pemodelan topik atau *topic modelling* dilakukan oleh Asmussen pada tahun 2019 melakukan penelitian pada sekumpulan literatur

menggunakan pemodelan topik. Langkah dari penelitian ini adalah pengumpulan data, *pre-processing*, pemodelan topik, dan *post-processing*. Penulis ini menggunakan metode *Latent Dirichlet Allocation* (LDA). Data dalam penelitian ini adalah literature yang berjumlah 650 judul. Hasilnya adalah, terbentuk beberapa topik beserta distribusi kata-kata dari setiap topik [15].

Penelitian data teks pada media sosial dilakukan oleh Wu pada tahun 2015 yang menganalisa bagaimana media sosial, khususnya *Twitter* dapat menjadi sarana pembantu dalam bisnis. Penelitian ini menggunakan tools *Leximancer* dengan tujuan untuk mengetahui bagaimana perusahaan menggunakan media sosial untuk mengembangkan bisnisnya [16][8]. Hasilnya adalah tema apa yang sering dibicarakan di *Twitter* milik *Costco* dan *Walmart* dan bagaimana sentiment masyarakat tentang produk yang dimiliki oleh kedua retailer tersebut. Mostafa pada tahun 2020 juga melakukan analisis teks pada *Twitter* untuk mengetahui topik percakapan mengenai makanan halal. Dalam penelitian ini dilakukan analisis pemetaan *tweet* berdasarkan lokasi, mencari jaringan kata berdasarkan hubungan antar-*cluster*, sentiment, dan topik yang terbentuk berdasarkan probabilitas kata. Metode yang diterapkan dalam penelitian ini adalah *Lexicon Sentiment Analysis*, *Topic Modelling*, *Semantic Network Analysis* [17].

Review konsumen juga menjadi objek penelitian bagi Chen untuk melakukan analisis pendapat konsumen terhadap restoran makanan cepat saji [1]. Penelitian ini menerapkan metode *feature selection* menggunakan LASSO yang merupakan algoritma yang terdapat pada *MATLAB R2017*. Data teks yang dianalisis diperoleh dari *Trustpilot.com* dan website resmi milik restoran. Tujuan utama dari

makalah ini adalah mengetahui faktor yang mempengaruhi kepuasan pelanggan. Hasilnya adalah kata beserta hasil perhitungan *tf-idf* yang kemudian dilakukan pengelompokan kata secara manual untuk mengambil kesimpulan tentang hal yang menjadi faktor penentu kepuasan pelanggan.

Analisis jaringan semantik dilakukan oleh Park pada tahun 2019 untuk menganalisis faktor-faktor yang mempengaruhi pertukaran budaya antara China dan Korea [18]. Data yang digunakan sebagai bahan analisis adalah kata pencarian yang terkait dengan pertukaran budaya pada situs *Naver*, *Daum*, dan *Baidu*. Hasil dari analisis jaringan semantik pada portal Korea adalah “*Korean Wave*”, sedangkan di China pertukaran budaya lebih pada bidang pemerintahan. Jung pada tahun 2020 melakukan analisis semantik pada 1856 penelitian terkait *text mining* [19]. Hasil dari analisis tersebut adalah penggunaan metode *text mining* dari tahun 1980 sampai saat ini mengalami kenaikan dan sudah merambah ke berbagai bidang, seperti biologi, penambangan data, dan penelitian mengenai obat-obatan.

Dilihat dari Tabel 2.1, terdapat beberapa penelitian sebelumnya yang menerapkan konsep penambangan teks dengan menerapkan berbagai metode, seperti analisis sentiment, pemodelan topik, jaringan semantik, dan menggunakan *tools* penambangan teks yang menangani sekumpulan teks secara otomatis. Terdapat berbagai bidang juga yang dilakukan analisis, seperti restoran, makanan, dan retailer. Dalam penelitian yang akan penulis lakukan, penulis menerapkan analisis teks yang datanya diambil dari *Twitter* dan *TripAdvisor* yang membahas tentang restoran makanan cepat saji yaitu Restoran K dan Restoran M. Lalu akan dilakukan analisis teks menggunakan metode pemodelan topik yang menerapkan

algoritma *LDA* dan kemudian dilanjutkan dengan melakukan visualisasi hasil. Penggunaan kedua metode ini dikarenakan *topic modelling* akan menghasilkan sekelompok kata yang memiliki nilai probabilitas yang setara, dan visualisasi terhadap hasil tersebut akan menyajikan sebaran dari kata-kata tersebut. Dengan mengkombinasikan kedua metode ini, dapat terlihat jelas bagaimana sebaran kata dan bagaimana topik yang terbentuk dari kata-kata tersebut [10]. Dari rangkaian analisis tersebut akan menghasilkan kesimpulan dari hasil analisis perhitungan menggunakan metode *LDA* dikaitkan dengan faktor pendorong pelanggan untuk melakukan pembelian.

Tabel 2.1 Perbandingan Penelitian

Penulis	Judul Penelitian	Metode	Hasil
Carracedo, et al [12]	<i>Research lines on the impact of the COVID-19 pandemic on business. A text mining analysis</i>	<i>Clustering Chi-Squared</i>	Ditemukan kata-kata yang sering muncul dalam <i>paper</i> yang dianalisis dan dapat divisualisasikan menjadi tiga <i>cluster</i> topik utama, yaitu pengaruh pandemic terhadap manajemen kesehatan, perubahan dalam kebiasaan konsumsi masyarakat, dan pengaruh pandemic terhadap organisasi bisnis.
He, et al [16]	<i>Gaining Competitive Intelligence from Social Media Data: Evidence from Two Largest Retail Chains</i>	<i>Clustering dan Sentiment Analysis menggunakan Tools Leximancer</i>	Hasil dari analisis tersebut adalah bagaimana pandangan konsumen (positif, negative, netral)

	<i>in the World Introduction</i>		terhadap produk milik masing masing retailer dan <i>cluster</i> dari tema pembicaraan pada <i>Twitter</i> .
Mostafa [17]	<i>Global halal food discourse on social media: a text mining approach</i>	<i>Lexicon Sentiment Analysis, Topic Modelling, Semantic Network Analysis</i>	Dari pemetaan <i>tweet</i> , tidak ditemukan <i>tweet</i> tentang makanan halal di negara dengan warga mayoritas muslim karena hal itu tidak menjadi isu yang besar. Diperoleh juga jaringan kata berdasarkan hubungan antar- <i>cluster</i> , sentiment, dan topik yang terbentuk berdasarkan probabilitas kata.
Gan, et al [13]	<i>A Text Mining and Multidimensional Sentiment Analysis of Online Restaurant Reviews</i>	<i>Multidimensional Sentiment Analysis</i>	Skor rata-rata sentimen makanan, layanan, suasana, harga, dan konteks restoran semuanya positif, yang menunjukkan bahwa konsumen merasa positif tentang kelima atribut ini.
Chen, et al [1]	<i>Using a Text Mining Approach to Hear Voices of Customers from Social Media toward the Fast-Food Restaurant Industry</i>	<i>Feature selection menggunakan LASSO</i>	Daftar kata berdasarkan perhitungan <i>tf-idf</i> dan metode LASSO yang kemudian disimpulkan secara manual oleh peneliti sehingga ditemukan hal yang menjadi penentu faktor

			kepuasan pelanggan.
Park, et al [18]	<i>The main factors affecting cultural exchange between Korea and China: A semantic network analysis based on the cultural governance perspective</i>	<i>Semantic Network Analysis</i>	Hasil dari analisis jaringan semantic pada portal Korea adalah “ <i>Korean Wave</i> ”, sedangkan di China pertukaran budaya lebih pada bidang pemerintahan.
Pratidina, TM	PEMBANGKITAN PENGETAHUAN BERDASAR PEMODELAN TOPIK UNTUK MENINGKATKAN MINAT PEMBELIAN PADA RESTORAN CEPAT SAJI	<i>Topic modelling menggunakan Latent Dirichlet Allocation</i>	Kesimpulan dari hasil analisis perhitungan menggunakan metode <i>LDA</i> dan <i>Intertopic Distance Map</i> dikaitkan dengan faktor pendorong pelanggan untuk melakukan pembelian.