

Analisis Persepsi Konsumen Terhadap Kualitas Layanan Gojek Menggunakan *Sentiment Analysis* Dan *Topic Modeling* Berdasarkan *Deep Learning IndoBERT*

Consumer Perception Analysis Of Gojek Service Quality Using Sentiment Analysis And Topic Modeling Based On Deep Learning IndoBERT

Nofita Mahfudiyah¹, Andry Alamsyah²

¹ Manajemen Bisnis Telekomunikasi & Informatika, Fakultas Ekonomi Bisnis, Universitas Telkom, Indonesia, nofitamahfudiyah@student.telkomuniversity.ac.id

² Manajemen Bisnis Telekomunikasi & Informatika, Fakultas Ekonomi Bisnis, Universitas Telkom, Indonesia, andrya@telkomuniversity.ac.id

Abstract

The development of smartphone-based applications in Indonesia is increasing in number thanks to the growing digital technology. This provides an opportunity for start-ups to come up with new innovations and compete to create output in the form of applications, one of which is the Gojek application, which is known as a pioneer in this sector. This study aims to analyze consumer perceptions of Gojek. This is intended as an effort to see the value of Gojek's service quality based on consumer tweets from Twitter social media. In addition, to find out the discussion of topics that are widely discussed by consumers. The discussion of these topics is used to determine the dimensions of improvements that need to be made by Gojek. The method used is deep learning which is one of the branches of Machine Learning (ML) using Deep Neural Network to solve problems. To achieve good results, a transfer learning approach such as IndoBERT is needed. This study uses IndoBERT in sentiment analysis and topic modeling analysis using BERTopic. The results showed 96% accuracy in sentiment analysis using IndoBERT. The proportion of sentiment is dominated by negative sentiment by 56% and the dimensions of mobile customer service and billing are the highest topics in negative sentiment. So that the results of the analysis in this study can provide benefits to the company to determine consumer perceptions and dimensions of service quality that need to be improved and developed by the company.

Keywords: consumer perception, Gojek, IndoBERT, sentiment analysis, topic modeling.

Abstrak

Perkembangan aplikasi berbasis smartphone di Indonesia semakin banyak jumlahnya berkat teknologi digital yang semakin berkembang. Hal ini memberi peluang kepada *start-up* untuk memunculkan inovasi baru dan bersaing menciptakan output berupa aplikasi, salah satunya di sektor transportasi muncul aplikasi Gojek yang dikenal sebagai pelopor di sektor ini. Penelitian ini bertujuan untuk menganalisis persepsi konsumen Gojek. Hal tersebut ditujukan sebagai upaya untuk melihat nilai kualitas layanan Gojek berdasarkan *uit-an* konsumen dari media sosial Twitter. Selain itu, untuk mengetahui pembahasan topik yang ramai diperbincangkan oleh konsumen. Pembahasan topik tersebut yang digunakan untuk mengetahui dimensi perbaikan yang perlu dilakukan oleh Gojek. Metode yang digunakan yaitu *deep learning* yang merupakan salah satu cabang *Machine Learning* (ML) menggunakan *Deep Neural Network* untuk menyelesaikan permasalahan. Untuk mencapai hasil yang baik, maka dibutuhkan sebuah pendekatan *transfer learning* seperti *IndoBERT*. Penelitian ini menggunakan *IndoBERT* dalam analisis sentimen dan analisis *topic modelling* menggunakan *BERTopic*. Hasil penelitian menunjukkan akurasi 96% pada analisis sentimen menggunakan *IndoBERT*. Proporsi sentimen didominasi oleh sentimen negatif sebesar 56% dan dimensi *mobile customer service and billing* menjadi topik tertinggi dalam sentimen negatif. Sehingga hasil analisis dalam penelitian ini dapat memberikan manfaat kepada perusahaan untuk mengetahui persepsi konsumen dan dimensi kualitas layanan yang perlu diperbaiki dan dikembangkan oleh perusahaan.

Kata kunci : analisis sentimen, Gojek, *topic modelling*, *IndoBERT*, persepsi konsumen

I. PENDAHULUAN

Perkembangan aplikasi berbasis *smartphone* di Indonesia semakin banyak jumlahnya, dibuktikan dengan riset dari *We Are Social* yang menunjukkan bahwa, jumlah total aplikasi yang diunduh pada tahun 2020 yaitu 6,32 miliar [1]. Hal ini memberi peluang kepada *start-up* untuk memunculkan inovasi baru dalam bersaing menciptakan aplikasi. Persaingan inovasi bisa terjadi di berbagai sektor, salah satunya di sektor transportasi dengan munculnya aplikasi layanan transportasi *online*. Gojek merupakan aplikasi layanan ojek *online* berbasis *smartphone* untuk menghubungkan antara pengguna jasa dan pengemudi. Berdasarkan siaran *pers* Gojek, aplikasi gojek telah diunduh lebih dari 142 juta kali dengan lebih dari 2 juta mitra pengemudi [2]. Aplikasi Gojek menduduki posisi pertama di *App Store* dan *Play Store* sebagai *top free apps* pada kategori *travel & local* [3]. Hasil survei dari Komunitas Konsumen Indonesia menyebutkan bahwa, dalam memilih layanan ojek *online*, 36 persen masyarakat lebih banyak menggunakan aplikasi Gojek daripada Grab yang memiliki persentase sebesar 32 persen. Sisanya yaitu 32 persen menggunakan Gojek dan Grab [4]. Hal tersebut memperlihatkan bahwa kehadiran aplikasi Gojek disambut baik oleh masyarakat.

Gojek yang dikenal sebagai perusahaan pelopor di sektor transportasi perlu menjaga, memperhatikan serta meningkatkan kualitas pelayanan dan menciptakan kepuasan konsumen, serta memiliki kemampuan serta kekuatan untuk bersaing dalam menguasai pasar [5]. Terdapat penelitian terdahulu yang menggunakan Gojek sebagai objek penelitiannya. Analisis sentimen yang diperoleh Gojek sebesar 40% sentimen negatif dan 60% sentimen positif, sehingga persepsi konsumen terhadap Gojek dinilai cukup baik [6]. Hasil untuk *topic modelling* yaitu pembahasan tentang manfaat setelah menggunakan layanan Gojek, layanan Gojek di daerah Bandung dan Banyuwangi, lalu 3 topik lainnya pembicaraan mengenai pengalaman di *android* dan *ios* dalam melakukan pemesanan dan pembatalan. Namun, penelitian dari [7] yang meneliti sentimen dari sisi ulasan di *Playstore* menghasilkan sentimen negatif dengan peluang kemunculan kata sebesar 1,3125. Berdasarkan literatur yang telah dikumpulkan peneliti, yaitu layanan Gojek yang terintegrasi dalam 1 aplikasi dan menjadi salah satu *unicorn* Indonesia [8]. Maka dari itu, agar Gojek dapat mempertahankan posisi dan mengembangkan inovasi, serta memperhatikan sentimen pengguna terhadap keberlangsungan perusahaan, penting untuk melakukan penelitian tentang persepsi konsumen terhadap layanan Gojek. Tujuan yang dicapai dalam penelitian ini adalah menganalisis sentimen konsumen terhadap kualitas layanan Gojek, menganalisis pembahasan topik yang ramai diperbincangkan oleh konsumen terhadap kualitas layanan Gojek berdasarkan media sosial Twitter, dan menganalisis dimensi perbaikan yang perlu dilakukan oleh Gojek dari hasil sentimen dan *topic modelling*.

II. METODOLOGI PENELITIAN

Pengumpulan data dilakukan menggunakan *Google Collaboratory* untuk melakukan *crawling* data pada media sosial Twitter. Setelah memperoleh data, data disimpan dalam format *.csv* di *Microsoft Excel 365*. Sebelum data diolah, peneliti melakukan pembersihan data yang disebut dengan *pre-processing text* yang sudah terkumpul di *Microsoft Excel 365*. Proses ini berfungsi untuk menghindari data yang berulang, tidak konsisten, tidak sempurna. Tahapan pertama yaitu melakukan *filtering* terhadap *duplicate tweets* dengan tujuan menghapus *tweets* yang memiliki isi yang sama sehingga tidak terjadi ketidakseimbangan data. Tahap kedua yaitu melakukan *cleaning data* yaitu dengan menghapus karakter selain huruf karena karakter tersebut dianggap *delimiter*. Selain itu, menghapus tautan, *mentions* dan *hashtags*.

Setelah data melalui proses *pre-processing text*, langkah berikutnya adalah *sentiment analysis*. Pada penelitian ini, proses dalam *sentiment analysis* menggunakan metode *IndoBERT*. Data yang telah melalui *pre-processing* kemudian diberi label dan diberi nama data training berdasarkan sentimen konsumen terhadap kualitas layanan Gojek, seperti tabel 2.1.

Tabel 2. 1 Contoh Pelabelan *Data Training*

Teks	Sentimen
Driver nya ramah dan berkualitas	Positif
Nggak sesuai harapan	Negatif

Saya baru install aplikasi gojek	Neutral
----------------------------------	---------

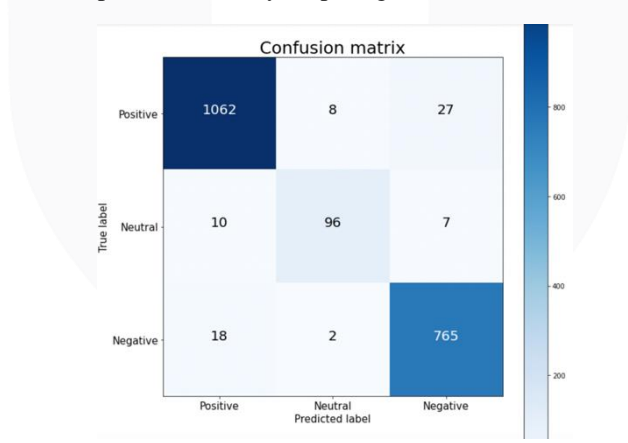
Data yang sudah diberi label, dibagi menjadi 3 proporsi, yaitu *data training*, *data testing* dan *data validation*. Pada penelitian ini, proporsi tersebut memiliki 3 varian, yaitu 80:10:10, 70:15:15 dan 60:20:20. Dari ketiga varian proporsi *dataset*, akan dilihat hasil akurasi *data validation* dan *data testing* tertinggi. Proporsi *dataset* yang memiliki akurasi *data validation* dan *data testing* tertinggi digunakan untuk *data testing* selanjutnya dan *topic modelling*.

Tahap selanjutnya adalah tahap *Topic Modelling*. Pada tahap ini, data yang sudah melalui tahap *sentiment analysis* akan diolah dengan *topic modelling*. Pengolahan ini berdasarkan klasifikasi dari sentimen positif, *neutral* dan negatif. Sehingga bisa dilihat topik yang ramai dibahas berdasarkan klasifikasi dari *sentiment analysis*. Algoritma yang digunakan untuk *topic modelling* merupakan *BERTopic*. Langkah dalam menggunakan *BERTopic* yaitu *dataset* harus berupa *string* dan mengatur data kedalam metode *fit_transform*. Terdapat dua hasil, yaitu hasil untuk topik dan hasil probabilitas. Hasil dari topik merupakan urutan topik yang berkaitan, sedangkan hasil dari probabilitas menunjukkan kemungkinan kalimat tersebut dikelompokkan ke dalam salah satu topik yang mungkin akan terjadi.

III. HASIL DAN PEMBAHASAN

Data yang akan diolah dalam penelitian ini berasal dari sosial media Twitter. Data berupa *tweet* dari pengguna Gojek yang diposting dengan *tagline* akun *gojekindonesia* dan *gojek* dalam rentang waktu 3 bulan, terhitung mulai tanggal 1 Oktober 2021 sampai 31 Desember 2021. Karakteristik data yang diperoleh berbahasa Indonesia dan termasuk ke dalam data tidak terstruktur. Proses pengambilan data dari Twitter dilakukan dengan metode *crawling*. *Crawling* dilakukan dengan menggunakan kata kunci “gojek” dan “gojekindonesia” sesuai dengan akun *official* Gojek Twitter. Jumlah data yang berhasil didapatkan sebesar 32895 Data. Data tersebut merupakan data mentah yang selanjutnya diproses ke tahap selanjutnya agar data bersih, rapi dan siap untuk diolah serta dianalisis.

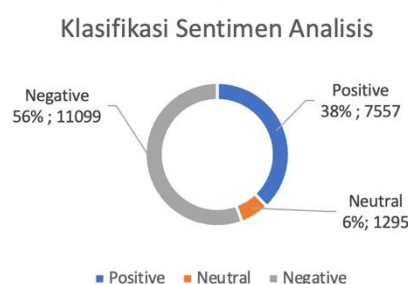
Pembagian *dataset* kedalam 3 proporsi, mendapatkan hasil akurasi tertinggi pada proporsi 80:10:10 dengan hasil akurasi *training* sebesar 98% dan akurasi *validation* sebesar 82%. Pada hasil akurasi *testing*, mendapatkan akurasi sebesar 96%. Selain dari hasil akurasi, peneliti menggunakan *confusion matrix* untuk mengukur sejauh mana model *IndoBERT* melakukan prediksi. Hasilnya seperti gambar 3.1 dibawah ini.



Gambar 3. 1 Confusion Matrix

Sumber: data olahan peneliti, 2022

Berdasarkan gambar 3.1 dapat dilihat bahwa model *IndoBERT* mampu mengklasifikasikan dengan benar label *positive* sebanyak 1062 dan salah dalam memprediksi label lain sebanyak 35. Label *neutral* berhasil di prediksi dengan benar sebanyak 96 dan salah dalam memprediksi label lain sebanyak 17 data. Label *negative* berhasil diprediksi dengan benar sebanyak 765 dan salah dalam memprediksi label lain sebanyak 20.

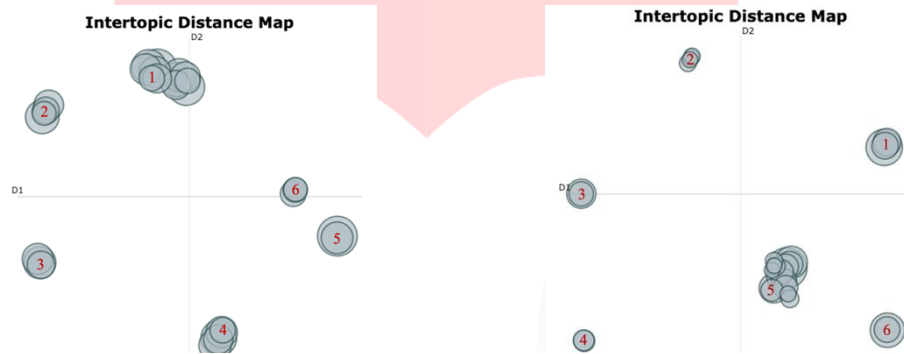


Gambar 3. 2 Proporsi Sentimen Keseluruhan

Sumber: data olahan peneliti, 2022

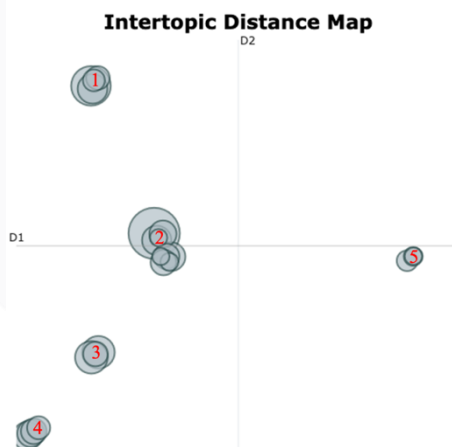
Pada gambar 4.3 menunjukkan bahwa sentimen negatif memiliki proporsi yang paling tinggi. Dari 100%, *sentiment negative* Memiliki presentase 56% atau 11099 Data. Sebesar 38% dimiliki oleh *sentiment positive* dengan jumlah 7557 data. Berdasarkan jumlah proporsi tersebut, dapat ditarik kesimpulan bahwa konsumen dari Gojek merasa kecewa atau kurang puas terhadap kualitas layanan yang diberikan oleh Gojek.

Selanjutnya dilakukan analisis *topic modelling* pada tiap sentimen. Data yang digunakan dalam analisis *topic modelling* adalah data yang sudah melalui tahapan *pre-processing* dan sudah berlabel. Peneliti membagi proses *topic modelling* untuk setiap sentimen (*positive, negative, neutral*). Sehingga bisa diketahui bahwa dalam setiap sentimen, topik apa saja yang paling banyak dibahas dengan melihat dari frekuensi kata yang sering muncul.



Gambar 3. 3 Clustering Sentimen Positif

Gambar 3. 4 Clustering Sentimen Negatif



Gambar 3. 5 Clustering Sentimen Neutral

Sumber: data olahan peneliti, 2022

Hasil *clustering topic modelling* sentimen positif dapat dilihat dalam gambar 3.4. *Topic modelling* dalam sentimen positif paling banyak membahas tentang dimensi *mobile convenience & reliability*. Pembahasannya adalah tentang kenyamanan yang diterima oleh konsumen. Kenyamanan tersebut merupakan hasil dari aplikasi yang praktis, cepat dan memudahkan atau mudah untuk diakses. Dengan adanya aplikasi ini, konsumen juga dimudahkan dalam mobilitas keseharian dimasa pppm, dan mengandalkan Gojek dalam kesehariannya. Hal ini membuktikan bahwa konsumen menerima keuntungan disisi mobilitas dan aksesibilitas yang diberikan oleh Gojek.

Mobilitas dan aksesibilitas merupakan indikator dari dimensi *mobile convenience & reliability*. Dimensi ini memiliki frekuensi kemunculan 1950.

Topic modelling dalam sentimen negatif paling banyak membahas tentang dimensi *mobile customer service and billing*. Pembahasan dalam cluster ini adalah keluhan konsumen terhadap sistem pembayaran yang ada dalam aplikasi, terutama *gopay* dan *gopaylater* yang payah. *Feedback* yang diberikan oleh konsumen berupa *voucher* atau promo yang tidak bisa dipakai, kegagalan dalam *top up gopay* dan pembayaran *gopaylater*, gagal dalam pembelian pulsa yang mengakibatkan pemblokiran. Selain itu, konsumen sudah mengajukan ke *customer service*, namun belum ada respon baik. Pembahasan tersebut memuat indikator yang ada dalam dimensi *mobile customer service and billing*, yaitu mekanisme *feedback*, ketersediaan bantuan *online* atau *telephone* dan mekanisme pembayaran. Namun, karena pembahasan ini berupa pembahasan sentimen negatif, berarti analisis memiliki nilai negatif dalam setiap indikatornya. Dimensi ini memiliki frekuensi kemunculan 2569. Hasil *clustering topic modelling* sentimen negatif dapat dilihat dalam gambar 3.5.

Gambar 3.6. merupakan hasil *clustering topic modelling* sentimen *neutral*. *Topic modelling* dalam sentimen *neutral* paling banyak membahas tentang dimensi *mobile security and privacy*. Pembahasannya adalah tentang harapan yang dimiliki konsumen terhadap proses transaksi dalam *gopay* dan *gopaylater*. Konsumen memberikan masukan untuk mengganti metode dari *otp* menjadi *link* yang dikirim *via email*. Selain itu, dalam proses *upgrade* dan transaksi *gopay* ataupun *gopaylater*, membutuhkan informasi mengenai *ktp* dan *bank* dari konsumen. Hal ini diharapkan Gojek memberikan keamanan ketat terhadap transaksi keuangan yang dilakukan oleh konsumen serta kemanana informasi pribadi dari konsumen. Dimensi ini memiliki frekuensi kemunculan 505.

IV. KESIMPULAN

Hasil dari penelitian yang sudah dilakukan, terdapat tiga kesimpulan yang diperoleh. Kesimpulan pertama terkait pemodelan *IndoBERT*. Hasil pemodelan menggunakan *IndoBERT* menghasilkan akurasi 96% dengan rasio *dataset* 80:10:10. Akurasi *precision* yang didapat sebesar 93% yang berarti bahwa model ini baik dalam ketepatan memprediksi kebenaran dibandingkan dengan keseluruhan hasil prediksi. Selain itu, akurasi yang dihasilkan *recall* menghasilkan 95% yang berarti bahwa *IndoBERT* dapat mengembalikan seluruh informasi mengenai nilai prediksi pada tingkatan yang baik. Berdasarkan keseluruhan hasil dari pemodelan *IndoBERT*, bisa disimpulkan bahwa *IndoBERT* mampu mengklasifikasikan dan memprediksi dengan baik karena semua akurasi diatas 90%.

Selanjutnya, kesimpulan kedua tentang hasil klasifikasi analisis sentimen. Hasil dari data ini menunjukkan bahwa sentimen negatif mendominasi sebesar 56%. Sementara sentimen positif sebesar 38% dan *neutral* sebesar 6%. Sentimen negatif mencakup 11099 data, sentimen positif mencakup 7557 data dan sentimen *neutral* mencakup 1295 data. Berdasarkan hasil klasifikasi sentimen, konsumen Gojek lebih banyak memiliki persepsi yang kurang baik. Perasaan kecewa dan kurang puas terhadap kualitas layanan yang diberikan oleh Gojek.

Kesimpulan ketiga tentang hasil *topic modelling*. *Topic modelling* menghasilkan *cluster* yang baik, karena bisa diidentifikasi dan dianalisis berdasarkan dimensi kualitas layanan. Pada sentimen positif, dimensi *mobile convenience and reliability* memiliki frekuensi tertinggi. Selanjutnya dimensi *mobile customer service and billing* dan *mobile system efficiency and availability*. Pada sentimen negatif, dimensi *mobile customer service and billing* memiliki frekuensi tertinggi. Selanjutnya dimensi *mobile system efficiency and availability*, *mobile convenience and availability* dan *safety and personnel*. Pada sentimen *neutral*, hanya ada dua dimensi, yaitu *mobile security and privacy* yang memiliki frekuensi tertinggi dan selanjutnya *mobile convenience and reliability*.

REFERENSI

- [1] Kemp. Simon, "DIGITAL 2021: INDONESIA," 2021. <https://datareportal.com/reports/digital-2021-indonesia> (accessed Nov. 15, 2021).
- [2] R. K. Hastuti, "Gojek Telah Diunduh 142 Juta Kali," 2019. <https://www.cnbcindonesia.com/tech/20190423171542-37-68367/gojek-telah-diunduh-142-juta-kali> (accessed Nov. 13, 2021).
- [3] L. A. Hamdani, "Gojek rebranding, dan kalahkan Grab soal pengguna aktif bulanan," 2019. <https://www.tek.id/tek/gojek-rebranding-dan-kalahkan-grab-soal-pengguna-aktif-bulanan-b1Xj29eZz> (accessed Nov. 16, 2021).

- [4] D. R. Kusuma, “Survei KKI: Masyarakat Lebih Pilih Pakai Gojek daripada Grab,” 2019. <https://kumparan.com/kumparanbisnis/survei-kki-masyarakat-lebih-pilih-pakai-gojek-daripada-grab-1raMrFhP1GL/1> (accessed Nov. 02, 2021).
- [5] E. Mogaji and I. Erkan, “Insight into consumer experience on UK train transportation services,” *Travel Behaviour and Society*, vol. 14, pp. 21–33, Jan. 2019, doi: 10.1016/j.tbs.2018.09.004.
- [6] C. Septiani Hudaya, H. Fakhurroja, and A. Alamsyah, “Analisis Persepsi Konsumen Terhadap Brand Go-jek Pada Media Sosial Twitter Menggunakan Metode Sentiment Analysis dan Topic Modelling,” *Cindy Septiani Hudaya*, vol. 1, no. 6, pp. 664–673, 2019.
- [7] S. Wahyu Handani, D. Intan Surya Saputra, Hasirun, R. Mega Arino, and G. Fiza Asyrofī Ramadhan, “Sentiment analysis for go-jek on google play store,” in *Journal of Physics: Conference Series*, Apr. 2019, vol. 1196, no. 1. doi: 10.1088/1742-6596/1196/1/012032.
- [8] D. J. Bayu, “Indonesia Kini Punya 7 Unicorn, Siapa Saja?,” 2021. <https://databoks.katadata.co.id/datapublish/2021/07/28/indonesia-kini-punya-7-unicorn-siapa-saja> (accessed Oct. 27, 2021).

