

## **1. PENDAHULUAN**

### **1.1 Latar Belakang**

Pertumbuhan eksploratif sektor *e-commerce* di Indonesia, dengan Tokopedia sebagai salah satu pemain kunci, telah menciptakan pasar digital yang sangat dinamis. Seiring dengan kemudahan transaksi, media sosial, khususnya Twitter (X), telah menjelma menjadi *real-time feedback channel* di mana pelanggan secara masif menyuarakan pengalaman mereka terhadap produk, promosi, dan layanan.

Volume data teks pada Twitter (X) sangat besar dan tidak terstruktur memuat “suara pelanggan” yang sangat berharga mengenai kinerja operasional perusahaan, namun tidak efisien untuk diproses secara manual. Menanggapi tantangan ini, terdapat solusi melalui dua teknik utama, yakni analisis sentimen dan pengelompokan topik (*Clustering*). Analisis sentimen memungkinkan klasifikasi cepat apakah suatu *tweet* bersifat positif, netral, atau negatif. *Clustering* secara otomatis mengidentifikasi isu-isu atau tren pembicaraan utama yang muncul. Oleh karena itu, penelitian ini berfokus pada implementasi teknik *Classification* dan *Clustering* pada data *tweet* spesifik pelanggan Tokopedia, dengan tujuan mengekstrak pemahaman yang akurat mengenai persepsi pelanggan dan tren topik pembicaraan untuk menjadi dasar rekomendasi strategis yang *actionable* bagi perusahaan *e-commerce* dalam meningkatkan kepuasan pelanggan dan efisiensi operasional.

### **1.2 Tujuan Penelitian**

Penelitian ini memiliki tujuan umum untuk menganalisis persepsi pelanggan dan mengidentifikasi tren topik pembicaraan terkait produk dan layanan Tokopedia di platform media sosial Twitter (X) melalui pendekatan *Natural Language Processing* (NLP). Tujuan ini dipecah menjadi tiga tujuan khusus yang spesifik dan terukur. Pertama, mengimplementasikan model klasifikasi teks untuk secara otomatis mengklasifikasikan setiap *tweet* ke dalam kategori sentimen: Positif, Netral, atau Negatif, serta mengevaluasi akurasi performa model. Kedua, menerapkan algoritma *clustering* pada korpus *tweet* untuk secara efektif mengidentifikasi dan mengelompokkan topik-topik utama pembicaraan pelanggan. Ketiga, menganalisis korelasi antara sentimen yang ditemukan dengan topik yang dikelompokkan untuk mengidentifikasi area kritis yang membutuhkan perbaikan dan memberikan rekomendasi strategis yang spesifik dan terperinci kepada perusahaan *e-commerce* Tokopedia.

## **2. METODOLOGI**

### **2.1 Sumber dan Pengumpulan Data**

Data yang digunakan dalam analisis jaringan ini bersumber secara eksklusif dari platform media sosial Twitter (X). Data diperoleh melalui teknik *web scraping/crawling* menggunakan API non-resmi yang berfokus pada kata kunci terkait layanan Tokopedia, seperti `tokopedia OR tokped, pengalaman AND tokopedia, dan tokopedia care`. Proses pengumpulan data mencakup rentang waktu selama satu tahun penuh, mulai dari Januari hingga Desember 2025. Secara total, dataset mentah yang dikumpulkan terdiri dari 2.113 *tweets* dan interaksi terkait isu tersebut.

### **2.2 Preprocessing Data**

Tahap *preprocessing data* dilakukan untuk mentransformasi data teks mentah dari *tweet* menjadi format yang bersih dan terstruktur, yang siap untuk dianalisis oleh model.

Proses ini dimulai dengan pembersihan *noise*, di mana karakter-karakter yang tidak relevan seperti URL, *mention*, dan *hashtag*, angka, dan tanda baca dihilangkan. Selanjutnya, dilakukan normalisasi teks dengan mengubah seluruh teks menjadi huruf kecil dan menggunakan kamus normalisasi yang telah disusun untuk mengonversi kata-kata *slang*, singkatan, dan *typo* ke dalam bentuk bakunya. Setelah teks bersih dan baku, proses tokenisasi diterapkan untuk memecah kalimat menjadi token-token atau kata individual. Terakhir, dilakukan *stopword removal* untuk membuang kata-kata umum dalam bahasa Indonesia yang tidak memiliki nilai informatif signifikan untuk analisis sentimen atau topik. Hasil akhir dari tahapan ini adalah korpus teks yang bersih dan siap untuk tahap *feature extraction* dengan jumlah data sebanyak 1697 data.

### 2.3 Model dan Teknik Analisis

Penelitian ini menggunakan tiga pendekatan komputasi yang berbeda untuk menganalisis data *tweet* secara komprehensif. Pertama, untuk Analisis Sentimen (*Classification*), model *Support Vector Machine* (SVM) dengan *kernel* linear (LinearSVC) akan diimplementasikan guna mengklasifikasikan *tweet* ke dalam sentimen Positif, Negatif, atau Netral. Teks *tweet* yang sudah bersih akan ditransformasi menjadi representasi numerik menggunakan TF-IDF (*Term Frequency-Inverse Document Frequency*), di mana model SVM dipilih karena keunggulannya dalam menangani data berdimensi tinggi. Kedua, untuk Pengelompokan Topik (*Clustering*), teknik *topic modeling Latent Dirichlet Allocation* (LDA) akan digunakan untuk mengidentifikasi topik-topik utama. LDA, sebagai model probabilistik yang mengasumsikan setiap dokumen adalah campuran beberapa topik, dipilih karena pendekatannya yang lebih realistik untuk data teks yang kompleks. Ketiga, Social Network Analysis (SNA) akan diterapkan dengan membangun directed graph di mana pengguna Twitter (X) adalah *Nodes* dan interaksi *mentions* adalah *Edges*.

### 2.4 Evaluasi Model

Setiap model akan dievaluasi menggunakan metrik yang sesuai dengan tujuannya. Kinerja Klasifikasi (SVM) akan diukur menggunakan Laporan Klasifikasi yang mencakup Akurasi, Presisi, *Recall*, dan F1-Score, dengan fokus khusus pada metrik Recall untuk kelas 'Negatif' guna memastikan kemampuan model mengidentifikasi sebanyak mungkin keluhan pelanggan secara efektif. Selanjutnya, kualitas *Clustering* (LDA) akan dievaluasi secara kuantitatif menggunakan Skor Koherensi, yang mengukur seberapa masuk akal dan relevan secara semantik kata-kata kunci dalam sebuah topik. Terakhir, Analisis Jejaring (SNA) akan dievaluasi dengan mengukur pengaruh akun menggunakan *In-Degree Centrality* (jumlah *mention* yang diterima) dan menilai kualitas deteksi komunitas dalam jaringan menggunakan Modularitas.

## 3. HASIL

### 3.1 Analisis Sentimen (*Classification*)

Implementasi klasifikasi sentimen diawali dengan pemuatan *dataset* final yang telah dilabeli, diikuti dengan penanganan data teks yang kosong. *Dataset* tersebut kemudian dibagi menjadi dua bagian, 80% sebagai data latih dan 20% sebagai data uji, dengan menerapkan untuk menjaga proporsi label sentimen di kedua set. Selanjutnya, teks pada data latih ditransformasi menjadi matriks numerik menggunakan TF-IDF *Vectorizer* (dibatasi 5000 fitur teratas), dan *vectorizer* yang sama diterapkan untuk mentransformasi data uji. Model

LinearSVG (SVM) kemudian dilatih menggunakan matriks TF-IDF dan label yang sebenarnya.

```
--- File Final Dimuat ---
Total baris: 1697

--- Distribusi Label Final ---
final_label
Netral      1230
Negatif     441
Positif     26
Name: count, dtype: int64

Memulai Proses Klasifikasi (SVM)...

Data dibagi: 1357 baris latih, 340 baris uji.

Melatih model SVM pada data final Anda...
Model SVM berhasil dilatih.

--- Laporan Klasifikasi (SVM) ---
Ini adalah evaluasi performa model Anda pada data uji:
      precision    recall   f1-score   support
Negatif       0.85     0.68     0.75      88
Netral        0.88     0.96     0.91     247
Positif        0.00     0.00     0.00       5

accuracy          0.87      340
macro avg        0.57     0.55     0.56     340
weighted avg     0.86     0.87     0.86     340
```

Gambar 1. Evaluasi Performa Model pada Data Uji

Hasil klasifikasi menunjukkan bahwa akurasi keseluruhan sebesar 87% telah dicapai. Namun, kinerja bervariasi antar kelas. Performa yang sangat kuat ditunjukkan dalam mengidentifikasi sentimen ‘Netral’ (F1-Score 0.91) dan kinerja yang cukup baik untuk sentimen ‘Negatif’ (F1-Score 0.75). Untuk kelas ‘Negatif’, *Recall* sebesar 0.68 dicapai, yang berarti 68% keluhan berhasil diteukan. Identifikasi sentimen ‘Positif’ (F1-Score 0.00) gagal dilakukan, yang disebabkan oleh jumlah sampel Positif yang sangat sedikit dalam *dataset* (data sangat imbalanced).

### 3.2 Clustering

Untuk mengidentifikasi topik utama dalam *tweet*, model *Latent Dirichlet Allocation* (LDA) diterapkan. Proses ini diawali dengan memuat *dataset* yang telah dilabeli, kemudian dilakukan pemfilteran data untuk memfokuskan *clustering* hanya pada *tweet* dengan sentimen ‘Negatif’, sehingga topik yang dihasilkan secara spesifik merepresentasikan keluhan pelanggan. Dari total 1967 data, teridentifikasi 441 data negatif yang kemudian ditransformasi menjadi matriks *Document-Term* menggunakan *CountVectorizer*. Selama proses ini, *stopword* kustom yang relevan dengan domain seperti ‘tokopedia’, ‘tokped’ dan lainnya dihilangkan, dan *vectorizer* diatur untuk mengabaikan kata-kata yang terlalu umum atau terlalu langka. Model *LatentDirichletAllocation* kemudian dilatih pada matriks ini dengan jumlah 5 topik untuk mengidentifikasi 5 keluhan utama. Terakhir, *library* ‘pyLDAvis’ digunakan untuk membuat visualisasi interaktif dari hasil model LDA, yang memetakan topik dan relevansi kata kunci.

--- Memuat Data ---

Data Negatif yang akan di-cluster: 441 baris

--- Kata Kunci Teratas Sesuai Urutan Visualisasi ---

Urutan di bawah ini sudah SAMA dengan urutan di visualisasi (Topik 1 = Lingkaran Terbesar)

Visualisasi Topik 1: mau | dana | kalo | pengalaman | jadi | beli | saldo | mohon | pengembalian | masuk

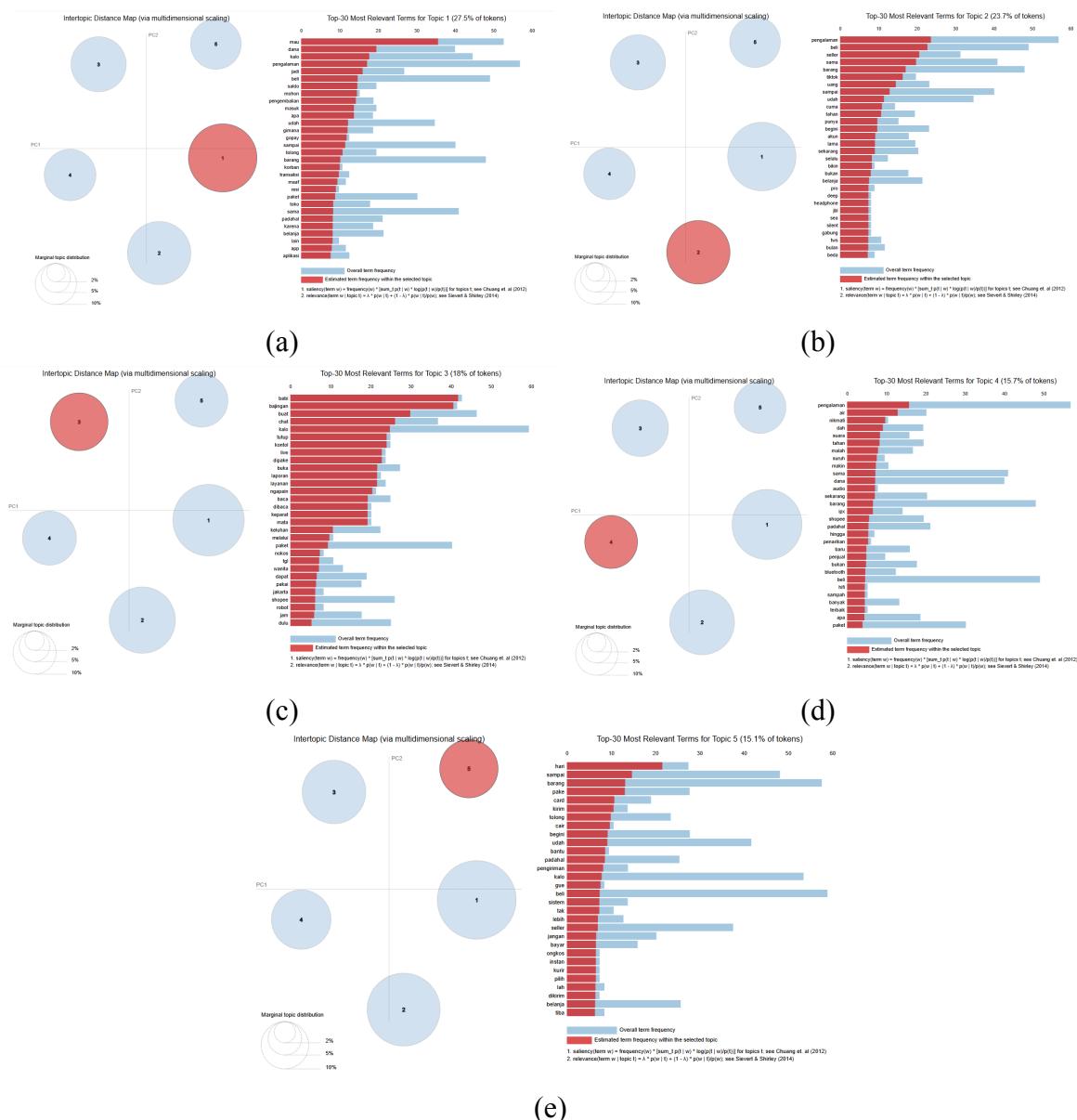
Visualisasi Topik 2: pengalaman | beli | seller | sama | barang | tiktok | uang | sampai | udah | cuma

Visualisasi Topik 3: babi | bajingan | buat | chat | kalo | tutup | kontol | live | dipake | buka

Visualisasi Topik 4: pengalaman | air | nikmati | dah | suara | tahan | malah | suruh | makin | sama

Visualisasi Topik 5: hari | sampai | barang | pake | card | kirim | tolong | cair | begini | udah

Gambar 2. 5 Cluster dengan 10 Kata Kunci Teratas



Gambar 3. Visualisasi Cluster dan Kata Kunci Teratas (a) Cluster 1, (b) Cluster 2, (c) Cluster 3, (d) Cluster 4, (e) Cluster 5

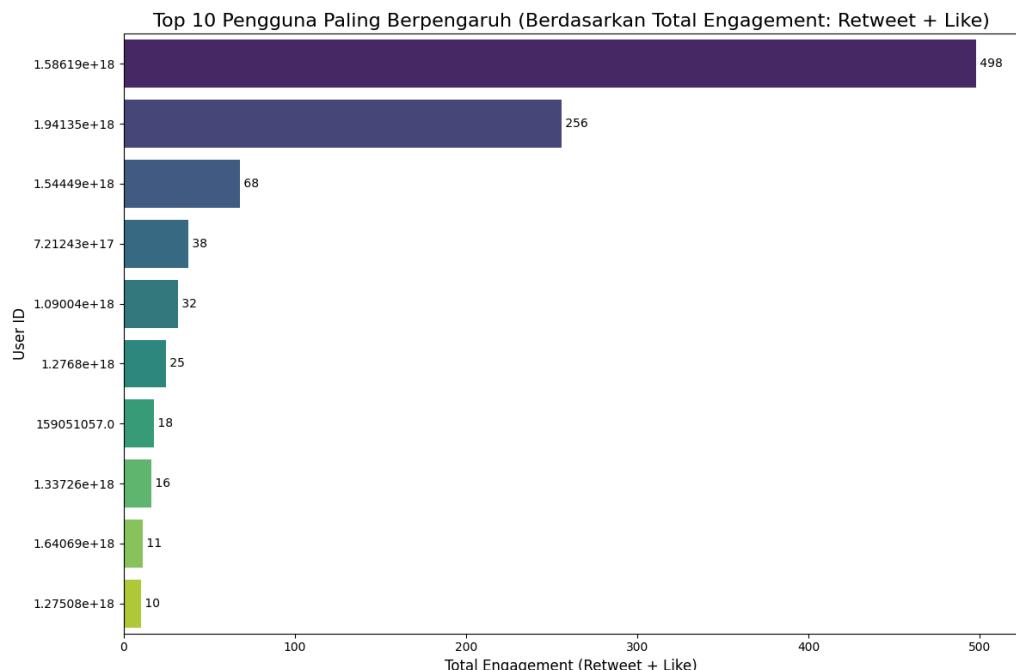
Output pada Gambar 2 menampilkan sepuluh kata kunci teratas yang mewakili masing-masing dari lima topik keluhan utama pengguna Tokopedia berdasarkan hasil analisis clustering terhadap data tweet bernada negatif. Berdasarkan kata kunci tersebut, setiap cluster dapat diinterpretasikan sebagai tema keluhan yang berbeda. Cluster 1 menunjukkan

permasalahan terkait dana dan saldo yang belum masuk serta proses pengembalian dana yang lambat, terlihat dari kemunculan kata seperti “*dana*”, “*saldo*”, “*pengembalian*”, dan “*masuk*”. *Cluster 2* berfokus pada keluhan mengenai pengalaman berbelanja dan interaksi dengan penjual (*seller*), termasuk perbandingan layanan dengan platform lain seperti TikTok *Shop*. *Cluster 3* berisi keluhan bernada emosional yang kuat dengan kata seperti “*babi*”, “*bajingan*”, dan “*chat*”, yang menunjukkan adanya gangguan pada fitur komunikasi seperti *chat* atau *live*. Sementara itu, *Cluster 4* menggambarkan ketidakpuasan terhadap kualitas atau kenyamanan produk, sedangkan *Cluster 5* berhubungan dengan durasi pengiriman dan pencairan dana yang lama.

Visualisasi pada Gambar 3 memberikan pemahaman yang lebih mendalam mengenai hubungan antar topik. *Section* pada bagian kiri menunjukkan proporsi dan jarak antar *Cluster*, di mana ukuran lingkaran merepresentasikan jumlah data dalam topik tersebut dan jarak antar lingkaran menunjukkan kesamaan semantik. Sedangkan, *section* sebelah kanan menampilkan diagram batang kata kunci yang paling relevan untuk topik yang dipilih.

### 3.3 Analisis Jejaring

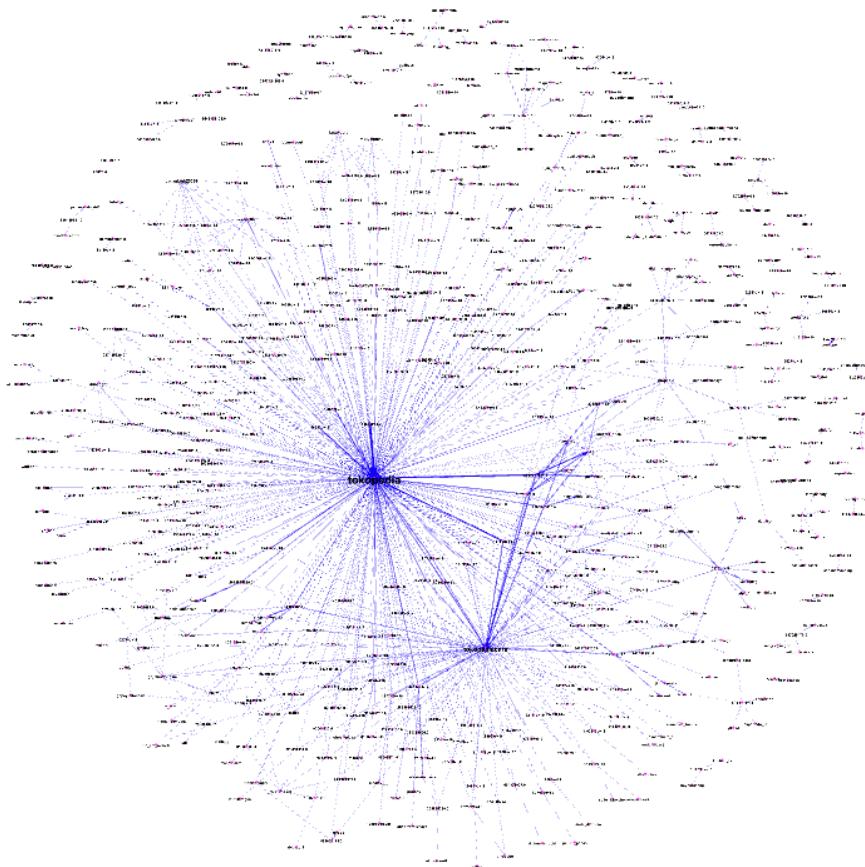
#### 3.3.1 Analisis Pengaruh *Engagement*



Gambar 4. 10 Pengguna Paling Berpengaruh Berdasarkan Total *Engagement*

Penjumlahan ‘retweet\_count’ dan ‘favorite\_count’ diagregasikan untuk setiap pengguna unik di dalam *dataset*. Sepuluh pengguna dengan total *engagement* tertinggi kemudian diekstraksi dan divisualisasikan seperti pada Gambar 4. Hasil visualisasi ini menunjukkan bahwa pengaruh *engagement* sangat tidak merata dan didominasi oleh segelintir akun. Dua pengguna teratas,  $1.58619e+18$  dengan 498 *engagement* dan  $1.94135e+18$  dengan 256 *engagement*, yang teridentifikasi terlibat dalam satu *thread* keluhan viral, memiliki jangkauan konten yang jauh melampaui pengguna lainnya.

### 3.3.2 Analisis Pengaruh Jaringan



Gambar 5. Graf Analisis Jejaring Sosial Tokopedia

Analisis pengaruh jaringan dilakukan dengan memvisualisasikan 1.238 interaksi *mention* ke dalam sebuah *directed graph*, seperti yang disajikan pada Gambar 5. Dalam graf yang terdiri dari 681 *nodes* dan 853 *edges* ini, ukuran *node* diatur berdasarkan *In-Degree Centrality* yakni jumlah *mention* yang diterima untuk mengidentifikasi akun paling berpengaruh atau yang menjadi pusat diskusi. Secara visual dan kuantitatif, graf didominasi oleh dua *node* sentral terbesar, yaitu @tokopedia dengan 312 *mention* dan @TokopediaCare dengan 119 *mention*, yang berfungsi sebagai pusat utama penerima keluhan.

Id	In-Degree ▼
@tokopedia	322
@tokopediacare	119
xkode06	15
dadankrb25038	14
shopeeid	10
gojekindonesia	10
jne_id	9
kutukobra	8
marukonahu	8
jntexpressid	7

Gambar 6. *In-Degree* Analisis Jejaring Sosial Tokopedia

Selain itu, analisis struktur komunitas menghasilkan Skor Modularitas 0,588. Skor di atas 0,5 mengindikasikan bahwa jaringan memiliki struktur komunitas yang kuat dan terdefinisi dengan baik.

### **Results:**

Modularity: 0.588  
Modularity with resolution: 0.588  
Number of Communities: 95

Gambar 7. Modularitas Analisis Jejaring Sosial Tokopedia

## **4. PEMBAHASAN**

### **4.1 Interpretasi Hasil**

Hasil analisis menunjukkan bahwa persepsi pelanggan terhadap Tokopedia di Twitter didominasi oleh sentimen Negatif dan Netral. Model klasifikasi SVM yang dikembangkan memiliki akurasi 87%, F1-Score Negatif 0,75 terbukti efektif sebagai “detektor keluhan”, namun gagal total (F1-Score 0,00) dalam mengidentifikasi sentimen Positif akibat data yang sangat tidak seimbang. Hal ini mengindikasikan bahwa Twitter (X) utamanya digunakan sebagai media komplain bagi pengguna. Analisis LDA pada 441 data negatif berhasil mengidentifikasi 5 *Cluster* keluhan utama, dengan topik paling dominan secara konsisten adalah masalah pengiriman & estimasi serta masalah *refund* & saldo. Temuan ini diperkuat oleh Analisis Jejaring (SNA), dimana @tokopedia dengan 312 *mention* dan @TokopediaCare dengan 119 *mention* menjadi *node* sentral dengan *In-Degree Centrality* tertinggi, membuktikan mereka adalah target utama dari keluhan-keluhan tersebut. Skor Modularitas 0,588 juga mengonfirmasi bahwa keluhan ini tidak terjadi secara terisolasi, melainkan dalam *Cluster* komunitas yang kuat dan terdefinisi dengan baik.

### **4.2 Insight**

*Key insight* dari *case study* ini adalah adanya perbedaan signifikan antara ‘pengaruh jaringan’ dan ‘pengaruh konten’. Graf analisis jejaring menunjukkan bahwa @tokopedia dan @TokopediaCare adalah *node* yang menjadi pusat percakapan. Namun, analisis *engagement* menunjukkan bahwa *tweet* paling viral justru berasal dari satu pengguna yang membahas satu kasus keluhan spesifik, yakni penipuan via *link* eksternal *seller*. Hal ini menyiratkan bahwa satu insiden layanan pelanggan yang sangat negatif dan viral dapat memiliki jangkauan yang jauh lebih besar dan lebih merusak reputasi daripada ratusan keluhan operasional harian yang ditujukan langsung ke akun layanan pelanggan.

### **4.3 Implikasi bagi Perusahaan**

Implikasi dari temuan ini bagi Tokopedia sangat jelas, Pertama, hasil LDA menunjukkan perlunya perbaikan operasional mendesak pada alur logistik tepatnya pada estimasi pengiriman serta kecepatan *refund* & *update saldo*, yang merupakan sumber kerugian utama pelanggan. Kedua, *insight* dari SNA menyarankan bahwa tim media sosial tidak bisa hanya fokus pada penanganan *volume* keluhan yang masuk ke @TokopediaCare. Perusahaan harus memiliki strategi peningkatan prioritas tinggi untuk secara proaktif mengidentifikasi dan menangani data negatif yang mulai viral, karena *thread* ini memiliki potensi kerusakan reputasi terbesar. Terakhir, rendahnya sentimen Positif yang terdeteksi mengindikasikan perlunya strategi pemasaran untuk mendorong pelanggan yang puas agar mau berbagi pengalaman positif mereka guna menyeimbangkan narasi publik.

## 5. PENUTUP

### 5.1 Kesimpulan

Penelitian ini berhasil menganalisis persepsi pelanggan dan tren topik Tokopedia di Twitter (X) melalui tiga pendekatan. Persepsi pelanggan didominasi oleh sentimen Negatif dan Netral, yang berhasil diklasifikasikan oleh model SVM dengan akurasi 87% (F1-Score 0,75 untuk Negatif). Analisis *Clustering* menggunakan LDA secara spesifik mengidentifikasi bahwa keluhan-keluhan ini berpusat pada dua masalah operasional utama, yakni logistik pengiriman dan proses *refund*. SNA memperkuat temuan ini dengan menunjukkan @tokopedia dan @TokopediaCare sebagai *node* sentral dan skor modularitas 0,588 yang tinggi mengonfirmasi bahwa diskusi keluhan ini terstruktur dalam komunitas-komunitas yang solid.

### 5.2 Rekomendasi

Berdasarkan keterbatasan penelitian ini, terdapat beberapa rekomendasi untuk penelitian selanjutnya. Pertama, untuk mengatasi kegagalan model dalam mendeteksi sentimen Positif, disarankan untuk menggunakan teknik data *augmentation* atau *resampling* pada kelas minoritas untuk menciptakan *dataset* latih yang lebih seimbang. Kedua, untuk meningkatkan akurasi klasifikasi dan pemodelan topik, penelitian selanjutnya dapat mengimplementasikan model *Deep Learning* berbasis *transformer*, yang memiliki pemahaman konteks dan semantik bahasa Indonesia yang lebih baik daripada SVM dan LDA. Terakhir, untuk optimasi LDA, disarankan untuk mengimplementasikan pencarian jumlah topik optimal secara otomatis dengan mengevaluasi dan memaksimalkan Skor Koherensi pada berbagai jumlah *Cluster*.

## LAMPIRAN

1. *Source Code (Crawling, Preprocessing, Classification, Clustering, visualisasi :*  
 [E-Commerce - Crawl data twitter > 2000 tweets - 19 Juli 2024.ipynb](#)
2. *Source File Visualisasi Analisis Jejaring menggunakan Gephi:* [SNA Case Study\\_Pyta Nur Chumairah.gephi](#)
3. *Kumpulan Dataset:*  [Dataset](#)