

Harap mengisi tabel ini, Tabel ini digunakan untuk keperluan komunikasi administrasi saja, saat publish akan dihapus oleh team editor.	
Nama author ke 1	
Nomor WA	
Prodi/Jurusan	
Perguruan Tinggi	

Analisis Sentimen Aplikasi Gojek Dengan Penerapan Model Random Forest Tree

Vito Faza Alfarizy, Putra Prakoso

Program Studi Sains Data, Universitas Nusa Mandiri
Jl. Margonda Raya No. 545, Pondok Cina Depok, Jawa Barat., Indonesia
Matanaqra03@gmail.com

ABSTRAK

Abstrak ditulis dalam **1 paragraf**, dan memuat (**Pendahuluan, Permasalahan, Tujuan, Metode, Hasil**). Artikel ini merupakan *template* Jurnal Skripsi Teknik Informatika dengan menggunakan MS-Word, dituliskan dalam **Bahasa Indonesia**. Banyak halaman antara **6 sampai 8** lembar dengan format A4-*two columns*. Halaman judul harus menyertakan judul yang spesifik, pengarang dan abstrak **maksimum 200 kata** pada awal makalah. Afiliasi dan alamat e-mail harus diberikan setelah nama pengarang. Penulisan Judul dengan menggunakan *Times New Roman* 12pt, **Bold**, *All caps*, Selain itu gunakan ukuran font 10pt.

Kata kunci : tuliskan maksimum 6 kata kunci di sini

1. PENDAHULUAN

Kemajuan teknologi digital yang sangat cepat telah mengubah cara masyarakat dalam berinteraksi dengan berbagai layanan. Aplikasi seluler kini menjadi bagian penting dalam kehidupan sehari-hari, salah satunya adalah aplikasi super app seperti Gojek. Sebagai salah satu platform teknologi terkemuka di Indonesia, Gojek menyediakan berbagai layanan mulai dari transportasi, pengantaran makanan, logistik, hingga pembayaran [1]. Setiap harinya, jutaan pengguna aktif menghasilkan volume data yang sangat besar, terutama berupa ulasan dan masukan di platform distribusi aplikasi seperti Google Play Store dan Apple App Store.

Ulasan dari pengguna merupakan sumber data teks yang kaya akan informasi mengenai opini, pengalaman, serta tingkat kepuasan terhadap layanan yang diberikan. Bagi perusahaan seperti Gojek, menganalisis data ini sangat penting untuk memahami kebutuhan pelanggan, menemukan kelemahan dalam layanan, serta membuat keputusan strategis guna meningkatkan kualitas produk [2]. Namun, karena jumlah ulasan yang sangat besar, analisis secara manual menjadi tidak praktis dan memerlukan banyak waktu. Oleh karena itu, pendekatan berbasis komputasi dibutuhkan untuk memproses data teks ini secara otomatis.

Di sinilah peran teknik Text Mining dan Analisis Sentimen menjadi sangat relevan. Analisis sentimen, atau opinion mining, merupakan cabang dari Natural Language Processing (NLP) yang bertujuan untuk mengidentifikasi, mengekstrak, serta mengklasifikasikan opini atau sentimen (positif, negatif, atau netral) dari data teks [3]. Melalui analisis

sentimen, perusahaan dapat memperoleh wawasan berharga secara cepat dari ribuan ulasan pengguna.

Sejumlah penelitian sebelumnya telah memanfaatkan algoritma machine learning untuk melakukan analisis sentimen terhadap ulasan aplikasi. Beberapa algoritma yang umum digunakan antara lain Naïve Bayes, Support Vector Machine (SVM), dan Decision Tree. Dalam penelitian ini, algoritma yang dipilih adalah Random Forest. Random Forest merupakan pengembangan dari metode decision tree, yang bekerja dengan membangun banyak pohon keputusan selama proses pelatihan, lalu menentukan hasil klasifikasi berdasarkan mayoritas suara dari pohon-pohon tersebut [4]. Algoritma ini dipilih karena keunggulannya dalam menangani data berdimensi tinggi, performa akurasi yang kompetitif, serta ketahanannya terhadap overfitting, menjadikannya sangat cocok untuk tugas klasifikasi teks [5].

Penelitian ini bertujuan untuk membangun model klasifikasi sentimen terhadap ulasan pengguna aplikasi Gojek dengan menggunakan algoritma Random Forest yang telah dioptimalkan melalui proses tuning. Model akan dilatih untuk mengklasifikasikan ulasan ke dalam tiga kategori sentimen: positif, negatif, dan netral. Evaluasi performa model akan dilakukan secara menyeluruh menggunakan metrik seperti akurasi, presisi, recall, dan F1-score, dengan analisis yang mendalam terhadap kemampuan model dalam mengidentifikasi masing-masing kategori sentimen. Hasil akhir dari penelitian ini diharapkan dapat menunjukkan efektivitas tinggi dari penggunaan Random Forest dalam klasifikasi sentimen serta memberikan wawasan yang berguna bagi pengembangan aplikasi dalam meningkatkan kualitas layanan.

2. TINJAUAN PUSTAKA

2.1. Text Mining

Text Mining, yang juga dikenal sebagai Text Data Mining atau Text Analytics, merupakan proses untuk mengekstraksi informasi serta pola yang berguna, bernilai tinggi, dan sebelumnya tidak diketahui dari data teks yang bersifat tidak terstruktur [6]. Di tengah perkembangan era digital, jumlah data berbasis teks semakin melimpah—mulai dari unggahan di media sosial, email, berita online, hingga ulasan pengguna seperti yang ditemukan pada aplikasi Gojek. Proses Text Mining memanfaatkan teknik dari linguistik komputasional, Natural Language Processing (NLP), serta algoritma machine learning untuk mengubah teks menjadi representasi numerik yang dapat dianalisis lebih lanjut guna mengungkap wawasan tersembunyi. Salah satu penerapan utama dari text mining adalah analisis sentimen.

2.2. Analisis Sentimen

Analisis sentimen, atau dikenal juga sebagai opinion mining, merupakan salah satu bidang studi yang fokus pada pengkajian opini, sikap, evaluasi, emosi, dan sentimen individu terhadap berbagai entitas seperti produk, layanan, organisasi, tokoh, isu, maupun topik tertentu [2]. Tujuan utama dari analisis ini adalah untuk mengidentifikasi polaritas dari suatu teks, yaitu apakah teks tersebut mengandung sentimen positif, negatif, atau netral. Dalam dunia bisnis, analisis sentimen terhadap ulasan pelanggan memiliki peran penting dalam riset pasar, pengawasan reputasi merek, serta peningkatan kualitas layanan.

Secara umum, terdapat dua pendekatan utama dalam analisis sentimen:

1. Pendekatan Berbasis Leksikon (Lexicon-based): Mengandalkan kamus sentimen yang berisi daftar kata-kata dengan nilai polaritas tertentu. Sentimen dari suatu teks dihitung dengan menjumlahkan skor polaritas dari kata-kata yang terdapat di dalamnya.
2. Pendekatan Berbasis Machine Learning: Menggunakan algoritma klasifikasi yang dilatih pada dataset berlabel untuk mempelajari pola dalam teks. Pendekatan ini cenderung lebih adaptif dan seringkali menghasilkan akurasi yang lebih tinggi karena mampu memahami konteks serta struktur kalimat yang kompleks.

Pendekatan yang digunakan dalam penelitian ini adalah pendekatan berbasis machine learning.

2.3. Algoritma Random Forest

Random Forest merupakan algoritma ensemble learning berbasis supervised learning yang terdiri dari kumpulan Pohon Keputusan (Decision Tree) [4]. Algoritma ini bekerja dengan membangun sejumlah pohon keputusan dari berbagai sub-sampel dataset, kemudian menggabungkan hasil prediksi dari masing-masing pohon—dengan cara rata-rata untuk regresi

atau voting/modus untuk klasifikasi—untuk menghasilkan prediksi akhir.

Pendekatan ini menggunakan metode Bootstrap Aggregating, atau Bagging, di mana setiap pohon dilatih menggunakan sampel acak dari dataset asli yang diambil dengan penggantian (with replacement). Selain itu, pada setiap titik pemisahan (split) dalam pohon, hanya sebagian fitur yang dipilih secara acak untuk dipertimbangkan.

2.4. Metrik Evaluasi

Untuk mengukur kinerja model klasifikasi, digunakan Confusion Matrix yang berisi informasi True Positive (TP), True Negative (TN), False Positive (FP), dan False Negative (FN). Dari matriks ini, beberapa metrik turunan dihitung:

- Akurasi (Accuracy): Persentase total prediksi yang benar.

$$(TP+TN) / (TP+FP+FN+TN)$$
- Presisi (Precision): Dari semua yang diprediksi sebagai kelas positif, berapa persen yang benar-benar positif. Berguna untuk mengukur tingkat kesalahan positif palsu.

$$(TP) / (TP + FP)$$
- Recall (Sensitivity): Dari semua data yang sebenarnya positif, berapa persen yang berhasil diprediksi dengan benar. Berguna untuk mengukur tingkat kesalahan negatif palsu.

$$TP / (TP + FN)$$
- F1-Score: Rata-rata harmonik dari Presisi dan Recall. Metrik ini sangat baik digunakan ketika terjadi ketidakseimbangan kelas dalam data.

$$(2 * Recall * Precision) / (Recall + Precision)$$

3. METODE PENELITIAN

3.1. Pengumpulan dan Pemilihan Data

Penelitian ini menggunakan data sekunder yang bersumber dari file CSV yang diunduh dari platform Kaggle. Meskipun data tersebut berasal dari ulasan pengguna di Play Store, pengumpulan data praktis dilakukan melalui Kaggle. Dari keseluruhan kolom yang tersedia, hanya tiga kolom yang relevan dan dipilih untuk analisis, yaitu userName, content (berisi teks ulasan), dan score (peringkat yang diberikan pengguna).

3.2. Pra-pemrosesan Data (Data Preprocessing)

Tahap pra-pemrosesan data bertujuan untuk membersihkan dan menyiapkan data teks mentah agar dapat diolah secara efektif oleh model. Langkah-langkah yang dilakukan adalah sebagai berikut:

- Pembersihan Data: Langkah pertama adalah menghapus baris data yang kosong (null) dan baris data yang terduplikasi untuk memastikan kualitas data. Setelah proses ini, tidak ada lagi data duplikat yang tersisa.

- Normalisasi Teks: Seluruh teks pada kolom content diubah menjadi huruf kecil (lowercase).
- Tokenisasi: Kalimat ulasan dipecah menjadi unit-unit kata (token) menggunakan fungsi `word_tokenize` dari library NLTK.
- Stopword Removal: Kata-kata umum yang tidak memiliki makna sentimen (contoh: "di", "dan", "yang") dihapus menggunakan daftar stopwords untuk Bahasa Indonesia dari library NLTK.
- Stemming: Setiap kata yang memiliki imbuhan diubah menjadi kata dasarnya. Proses ini dilakukan dengan menggunakan StemmerFactory dari library Sastrawi.

3.3. Pelabelan Sentimen

Dataset awal tidak memiliki label sentimen. Oleh karena itu, dilakukan proses pelabelan otomatis menggunakan library VADER (Valence Aware Dictionary and Sentiment Reasoner), yang merupakan metode berbasis leksikon yang dioptimalkan untuk teks informal seperti ulasan media sosial. Untuk meningkatkan akurasi pada konteks Bahasa Indonesia, sebuah kamus leksikon tambahan yang berisi kata-kata sentimen lokal (seperti 'kecewa', 'rugi', 'mahal') dibuat dan ditambahkan ke dalam kosakata VADER. Hasil dari proses ini adalah sebuah kolom baru bernama sentimen yang berisi label Positif, Negatif, atau Netral untuk setiap ulasan.

3.4. Ekstraksi Fitur dengan TF-IDF

Untuk mengubah data teks yang sudah bersih menjadi format numerik yang dapat diproses oleh algoritma machine learning, penelitian ini menggunakan metode TF-IDF (Term Frequency-Inverse Document Frequency). Metode ini akan memberikan bobot pada setiap kata, di mana kata yang sering muncul di satu ulasan namun jarang di ulasan lainnya akan dianggap lebih penting. Proses ini dilakukan dengan menggunakan `TfidfVectorizer` dari library Scikit-learn.

3.5. Pembangunan dan Evaluasi Model

Tahapan ini adalah inti dari proses klasifikasi sentimen. Dijelaskan sebagai berikut:

- Penyeimbangan Data: Distribusi data awal menunjukkan ketidakseimbangan kelas (Netral 49.3%, Positif 31.4%, Negatif 19.3%). Untuk menghindari bias model terhadap kelas mayoritas, data disampling ulang menggunakan teknik SMOTE (Synthetic Minority Over-sampling Technique) agar setiap kelas sentimen memiliki jumlah data yang merata.
- Pembagian Data: Dataset yang telah seimbang kemudian dibagi menjadi data latih (training) sebesar 80% dan data uji (testing) sebesar 20%.

- Pembangunan Model: Model klasifikasi dibangun menggunakan algoritma Random Forest Tree.
- Tuning Hyperparameter: Untuk mendapatkan performa model terbaik, dilakukan proses hyperparameter tuning menggunakan `RandomizedSearchCV`. Proses ini bertujuan mencari kombinasi parameter terbaik dari sebuah grid parameter yang telah ditentukan (`n_estimators`, `max_depth`, `min_samples_split`, `min_samples_leaf`). Pencarian ini menggunakan skema 5-fold cross-validation (`cv=5`) dengan accuracy sebagai metrik penilaian.
- Evaluasi Model: Model dengan parameter terbaik yang didapat dari proses tuning kemudian dievaluasi kinerjanya pada data uji. Metrik yang digunakan untuk evaluasi adalah Akurasi, Presisi, Recall, dan F1-Score yang disajikan dalam Classification Report.

3.6. Visualisasi Hasil Analisis

Untuk mendukung analisis, berbagai teknik visualisasi data digunakan. Histogram dan Funnel Chart dipakai untuk menggambarkan distribusi dan proporsi setiap kelas sentimen. Sementara itu, Word Cloud dan Donut Plot digunakan untuk menampilkan kata-kata yang paling sering muncul dan unik dari setiap kategori sentimen (Netral, Positif, dan Negatif).

4. HASIL DAN PEMBAHASAN

4.1. Analisis Data Eksploratif

Analisis data eksploratif dilakukan untuk memahami karakteristik dan distribusi dataset ulasan Gojek setelah melalui tahap pra-pemrosesan dan pelabelan.

4.1.1. Distribusi Sentimen

Setelah proses pelabelan, dilakukan perhitungan distribusi untuk setiap kelas sentimen. Hasilnya menunjukkan bahwa dataset tidak seimbang, dengan sentimen Netral sebagai kelas mayoritas.

- Distribusi Jumlah: Berdasarkan histogram, jumlah ulasan untuk setiap sentimen adalah:
 - Netral: 3.987 ulasan
 - Positif: 2.541 ulasan
 - Negatif: 1.562 ulasan
- Distribusi Proporsi: Visualisasi menggunakan Funnel Chart menunjukkan proporsi sebagai berikut:
 - Sentimen Netral mendominasi dengan 49.3% dari total ulasan, yang berarti hampir setengah dari ulasan bersifat netral.
 - Sentimen Positif menyusul dengan 31.4%, menunjukkan adanya jumlah pengguna yang puas yang cukup signifikan.
 - Sentimen Negatif mencakup 19.3% dari total ulasan. Meskipun minoritas, angka ini tetap signifikan

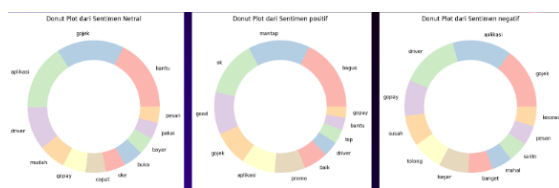
dan menjadi fokus penting untuk evaluasi layanan oleh Gojek. Ketidakseimbangan ini menjadi justifikasi untuk diterapkannya teknik resampling SMOTE pada tahap pemodelan untuk mencegah bias pada model.

4.1.2. Analisis Kata Kunci per Sentimen

Untuk memahami topik utama yang dibicarakan pada setiap kategori sentimen, dibuat visualisasi Word Cloud dan Donut Plot.



Gambar 4.1. Wordcloud sentimen netral, positif dan negatif



Gambar 4.2. Donut Plot sentimen netral, positif dan negatif

- Sentimen Positif: Kata-kata yang dominan adalah gojek, aplikasi, bagus, mantap, baik, promo, dan good. Hal ini mengindikasikan bahwa kepuasan pengguna umumnya terkait dengan kualitas aplikasi secara umum, kemudahan penggunaan, dan adanya program promosi.
- Sentimen Negatif: Kata kunci yang sering muncul antara lain driver, gojek, susah, kecewa, mahal, parah, sulit, dan gopay. Ini menunjukkan bahwa keluhan pengguna seringkali berpusat pada masalah dengan layanan driver, kesulitan dalam penggunaan aplikasi, kekecewaan, persepsi harga yang mahal, dan masalah terkait layanan pembayaran Gopay.
- Sentimen Netral: Kata-kata yang menonjol adalah gojek, bantu, aplikasi, mudah, cepat, driver, dan pesan. Kata-kata ini cenderung bersifat fungsional dan deskriptif, menjelaskan penggunaan aplikasi tanpa muatan emosi yang kuat.

4.2. Hasil Kinerja Model

Setelah melalui tahapan *tuning* dan pelatihan, model Random Forest dievaluasi kinerjanya pada data uji.

4.2.1. Hasil Tuning Hyperparameter

Proses *tuning* menggunakan RandomizedSearchCV berhasil menemukan kombinasi hyperparameter terbaik untuk model Random Forest. Parameter terbaik yang didapatkan adalah:

- `n_estimators`: 100
- `min_samples_split`: 5
- `min_samples_leaf`: 1
- `max_depth`: None

Konfigurasi ini menunjukkan bahwa model menggunakan 100 pohon keputusan, dengan batasan minimal 5 sampel untuk melakukan pemisahan node, dan membiarkan pohon tumbuh maksimal untuk menangkap pola data secara mendalam.

4.2.2. Kinerja Klasifikasi

Evaluasi model pada data uji menghasilkan *Classification Report* yang merangkum kinerja model secara keseluruhan dan untuk setiap kelas.

Metrik	Presisi	Recall	F1-Score	Support
Negatif	0.82	0.92	0.87	310
Netral	0.96	0.96	0.96	803
Positif	0.98	0.90	0.94	505
Akurasi			0.93	1618
Macro Avg	0.92	0.93	0.92	1618
Weighted Avg	0.94	0.93	0.94	1618

Tabel 4.1. Classification Report Model Random Forest (Tuned)

Berdasarkan hasil pada Tabel 4.1, dapat ditarik beberapa pembahasan penting:

- Kinerja Model Secara Keseluruhan Sangat Baik: Model mencapai akurasi keseluruhan sebesar 93%, yang menunjukkan bahwa model mampu mengklasifikasikan sentimen ulasan dengan sangat tepat. Nilai macro average dan weighted average untuk F1-Score (keduanya 0.92 dan 0.94) juga mengonfirmasi performa yang solid dan seimbang di semua kelas.
- Sangat Andal dalam Mengklasifikasikan Ulasan Netral dan Positif: Model menunjukkan kinerja yang luar biasa untuk kelas Netral dan Positif, dengan F1-Score masing-masing 0.96 dan 0.94. Terutama untuk kelas Positif, nilai presisi yang mencapai 0.98 menandakan bahwa ketika model memprediksi sebuah ulasan sebagai positif, prediksi tersebut hampir selalu benar.
- Efektif dalam Menangkap Ulasan Negatif.: Untuk ulasan Negatif, model mencapai F1-Score yang baik yaitu 0.87. Poin yang paling menonjol adalah nilai recall yang sangat tinggi (0.92). Ini berarti model sangat pandai dan sensitif dalam "menangkap" atau mengidentifikasi hampir semua ulasan yang sebenarnya bersifat negatif. Kemampuan ini

sangat krusial bagi perusahaan untuk dapat mendeteksi mayoritas keluhan pelanggan. Di sisi lain, nilai presisi untuk kelas negatif (0.82) sedikit lebih rendah dibandingkan kelas lain. Ini berarti ada sejumlah kecil ulasan yang sebenarnya bukan negatif, namun salah diklasifikasikan sebagai negatif oleh model. Terdapat trade-off di sini: model dioptimalkan untuk tidak melewatkan keluhan (recall tinggi), bahkan dengan risiko beberapa kali salah menduga (presisi lebih rendah). Dari sudut pandang bisnis, kondisi ini seringkali lebih disukai karena memastikan hampir semua umpan balik negatif dapat ditinjau.

5. KESIMPULAN DAN SARAN

5.1. Kesimpulan

Berdasarkan analisis dan evaluasi model yang telah dilakukan, dapat ditarik beberapa kesimpulan sebagai berikut:

- Model Random Forest Tree yang telah melalui proses tuning hyperparameter dan diterapkan pada data yang telah diseimbangkan dengan teknik SMOTE terbukti sangat efektif dan berhasil dalam melakukan klasifikasi sentimen pada ulasan pengguna aplikasi Gojek, dengan mencapai akurasi keseluruhan sebesar 93%.
- Model menunjukkan kinerja yang luar biasa dalam mengklasifikasikan ulasan dengan sentimen Netral dan Positif. Hal ini dibuktikan dengan perolehan nilai F1-score yang sangat tinggi, yaitu 0.96 untuk kelas Netral dan 0.94 untuk kelas Positif.
- Untuk ulasan dengan sentimen Negatif, model juga menunjukkan kinerja yang baik dengan F1-score sebesar 0.87. Keunggulan utama model pada kelas ini adalah nilai recall yang sangat tinggi (0.92). Ini mengindikasikan bahwa model memiliki kemampuan yang andal untuk menangkap dan mengidentifikasi sebagian besar keluhan atau umpan balik negatif dari pengguna, sebuah kapabilitas yang sangat berharga untuk sistem evaluasi layanan. Meskipun presisinya sedikit lebih rendah (0.82), kemampuan untuk meminimalkan ulasan negatif yang terlewatkan menjadi nilai lebih dari model ini.

5.2. Saran

Untuk pengembangan penelitian di masa depan, beberapa saran yang dapat dipertimbangkan adalah:

- Penggunaan Model yang Lebih Kompleks: Penelitian selanjutnya dapat mengeksplorasi penggunaan model deep learning seperti Long Short-Term Memory (LSTM), GRU, atau arsitektur Transformer (misalnya IndoBERT) yang memiliki kemampuan lebih

baik dalam memahami konteks dan semantik kalimat yang kompleks.

- Analisis Sentimen Berbasis Aspek (ABSA): Untuk mendapatkan wawasan yang lebih mendalam, penelitian dapat dikembangkan ke arah Aspect-Based Sentiment Analysis. Ini memungkinkan identifikasi sentimen terhadap aspek-aspek spesifik dari layanan Gojek, seperti "kualitas driver", "harga gofood", "antarmuka aplikasi", atau "kecepatan pembayaran".
- Pengembangan Leksikon Sentimen: Meskipun VADER telah digunakan, pengembangan leksikon sentimen yang lebih komprehensif untuk bahasa gaul, singkatan, dan istilah-istilah yang sering muncul dalam konteks ulasan aplikasi di Indonesia dapat lebih meningkatkan akurasi pada tahap pelabelan atau sebagai pendekatan hibrida.
- Eksplorasi Teknik Ekstraksi Fitur Lain: Selain TF-IDF, penelitian mendatang dapat mencoba teknik representasi teks yang lebih canggih seperti word embedding (Word2Vec, GloVe, atau FastText) untuk menangkap hubungan semantik antar kata dengan lebih baik.

DAFTAR PUSTAKA

- [1] I. S. Lundam and J. B. B. Darmawan, "Analisis Sentimen Ulasan Aplikasi Gojek Di Google Play Store Menggunakan Multinomial Naïve Bayes," *ReTII*, pp. 149–156, Nov. 2024.
- [2] "Sentiment Analysis 2nd Edition," Cambridge University Press & Assessment. Accessed: Jul. 11, 2025. [Online]. Available: <https://www.cambridge.org/universitypress/subjcts/computer-science/artificial-intelligence-and-natural-language-processing/sentiment-analysis-mining-opinions-sentiments-and-emotions-2nd-edition>
- [3] W. Medhat, A. Hassan, and H. Korashy, "Sentiment analysis algorithms and applications: A survey," *Ain Shams Engineering Journal*, vol. 5, no. 4, pp. 1093–1113, Dec. 2014, doi: 10.1016/j.asej.2014.04.011.
- [4] L. Breiman, "Random Forests," *Machine Learning*, vol. 45, no. 1, pp. 5–32, Oct. 2001, doi: 10.1023/A:1010933404324.
- [5] V. Kotu and B. Deshpande, *Data Science: Concepts and Practice*. Morgan Kaufmann, 2018.
- [6] "The Text Mining Handbook." Accessed: Jul. 11, 2025. [Online]. Available: <https://www.cambridge.org/core/books/text-mining-handbook/0634B1DF14259CB43FCCF28972AE4382>