

## ANALISIS SENTIMEN ALFAGIFT DENGAN MODEL NAIVE BAYES DAN PENYEIMBANGAN DATA SMOTE

Farhan Kurniansyah<sup>1</sup>, Rudi Kurniawan<sup>2</sup>, Bani Nurhakim<sup>3</sup>, Ade Rizki Rinaldi<sup>4</sup>.

Program Studi Teknik Informatika<sup>12</sup>  
Program Studi Manajemen Informatika<sup>3</sup>  
Program Studi Rekayasa Perangkat Lunak<sup>4</sup>

STMIK IKMI Cirebon  
<https://ikmi.ac.id/page/18/?lang=de>  
[farhankurniansyah1@gmail.com](mailto:farhankurniansyah1@gmail.com)

(\*) Corresponding Author : [farhankurniansyah1@gmail.com](mailto:farhankurniansyah1@gmail.com)  
Published : 30 Maret 2026

**Abstract**—The rapid growth of digital services has increased the need for sentiment analysis to understand user perceptions of mobile applications. Alfagift, a retail-based e-commerce application, receives thousands of user reviews that reflect both positive experiences and operational complaints. However, these reviews typically exhibit an imbalanced sentiment distribution, with positive reviews dominating negative ones. This imbalance causes classification models to become biased, reducing their ability to correctly identify minority sentiment classes. This study aims to examine the effect of the Synthetic Minority Oversampling Technique (SMOTE) on improving the performance of a Multinomial Naïve Bayes classifier in analyzing Alfagift user sentiment represented using TF-IDF weighting. The dataset consists of 3,500 Indonesian-language reviews collected through web scraping from the Google Play Store. All reviews were preprocessed using NLP techniques, including cleaning, case folding, normalization, tokenization, stopword removal, and stemming, followed by TF-IDF feature extraction. Model performance was evaluated using accuracy, precision, recall, and F1-score. The results show that before applying SMOTE, the model achieved an accuracy of 0.8937, while the recall for the positive class was only 0.77, indicating difficulty in identifying the minority class. After applying SMOTE to the training data, model performance improved, reaching an accuracy of 0.8967. The positive class also showed significant improvement, with precision increasing to 0.85, recall rising to 0.88, and the F1-score improving to 0.86. These improvements demonstrate that SMOTE successfully balanced the dataset, enabling the model to better detect minority sentiment without compromising performance on the majority class. The findings confirm that combining NLP preprocessing, TF-IDF feature extraction, Multinomial Naïve Bayes, and SMOTE is an effective approach to enhancing the accuracy and balance of sentiment classification for Alfagift user reviews.

**Keywords** : Sentiment Analysis, Naïve Bayes, SMOTE, TF-IDF, Alfagift, NLP.

**Abstrak**—Perkembangan layanan digital mendorong peningkatan pemanfaatan analisis sentimen untuk memahami opini pengguna terhadap aplikasi mobile. Alfagift, sebagai aplikasi e-commerce ritel modern, menerima ribuan ulasan pengguna yang mencerminkan pengalaman positif maupun keluhan operasional. Namun, ulasan tersebut umumnya memiliki distribusi sentimen yang tidak seimbang, di mana ulasan positif lebih dominan dibandingkan ulasan negatif. Kondisi ini menyebabkan model klasifikasi cenderung bias, sehingga performa dalam mendeteksi sentimen minoritas menjadi rendah. Penelitian ini bertujuan untuk menganalisis pengaruh penerapan teknik *Synthetic Minority Oversampling Technique* (SMOTE) terhadap peningkatan kinerja model *Multinomial Naïve Bayes* dalam klasifikasi sentimen ulasan Alfagift yang direpresentasikan menggunakan pembobotan TF-IDF. Dataset penelitian terdiri dari 3.500 ulasan berbahasa Indonesia yang diperoleh melalui *web scraping Google Play Store*. Seluruh data diproses melalui tahapan NLP, mencakup *cleaning, case folding, normalisasi, tokenisasi, stopword removal, dan stemming* sebelum dilakukan ekstraksi fitur TF-IDF. Model dievaluasi menggunakan metrik akurasi, *precision, recall, dan F1-score*. Hasil penelitian menunjukkan bahwa sebelum penerapan SMOTE, model menghasilkan akurasi sebesar 0,8937 dengan *recall* sentimen positif hanya 0,77, menandakan bahwa model kurang mampu mengenali kelas minoritas. Setelah SMOTE diterapkan pada data latih, performa model meningkat menjadi akurasi 0,8967, dengan perbaikan signifikan pada kelas positif: *precision* mencapai 0,85, *recall* meningkat menjadi 0,88, dan *F1-score* naik menjadi 0,86. Perbaikan tersebut mengindikasikan bahwa

teknik SMOTE berhasil menyeimbangkan distribusi data sehingga model lebih sensitif dalam mendeteksi sentimen minoritas tanpa mengorbankan stabilitas performa pada kelas mayoritas. Temuan ini membuktikan bahwa kombinasi preprocessing NLP, TF-IDF, *Multinomial Naïve Bayes*, dan SMOTE efektif digunakan untuk meningkatkan akurasi serta keseimbangan klasifikasi sentimen pada ulasan aplikasi Alfagift.

**Kata Kunci :** Analisis Sentimen, Naïve Bayes, SMOTE, TF-IDF, Alfagift, NLP.

## INTRODUCTION

Sentiment analysis merupakan bidang penting dalam kecerdasan buatan yang berfokus pada penggalian opini dan emosi dari teks pelanggan untuk mendukung pengambilan keputusan berbasis data di sektor e-commerce. Teknologi ini memungkinkan perusahaan mengolah ribuan ulasan pelanggan menjadi informasi strategis bagi peningkatan kualitas produk, perbaikan layanan, maupun penguatan strategi pemasaran (Bellar A.; Ballafkih M., 2024; Yang et al. 2020). Selain itu, analisis sentimen berkontribusi dalam pengembangan sistem rekomendasi dan pemetaan preferensi konsumen melalui identifikasi pola perilaku dan tren pasar secara lebih cepat.

Dalam praktiknya, analisis sentimen menghadapi tantangan ketidakseimbangan data (data imbalance), khususnya ketika proporsi ulasan positif jauh lebih dominan dibandingkan ulasan negatif. Situasi ini menyebabkan model cenderung bias terhadap kelas mayoritas dan menurunkan kemampuan dalam mengenali opini minoritas yang penting bagi evaluasi layanan. Tantangan ini semakin kompleks karena karakteristik data teks yang tidak terstruktur, bervariasi, dan sering mengandung bahasa informal atau kesalahan penulisan [3], [4]. Meskipun metode deep learning menawarkan performa yang kompetitif, teknik tersebut memerlukan dataset besar dan seimbang untuk mencapai akurasi optimal. Oleh sebab itu, algoritma yang lebih sederhana dan interpretatif seperti Naïve Bayes tetap relevan, terutama jika dikombinasikan dengan teknik oversampling seperti SMOTE (Kosolwattan aet al., 2023; Mukherjee & Khushi, 2021).

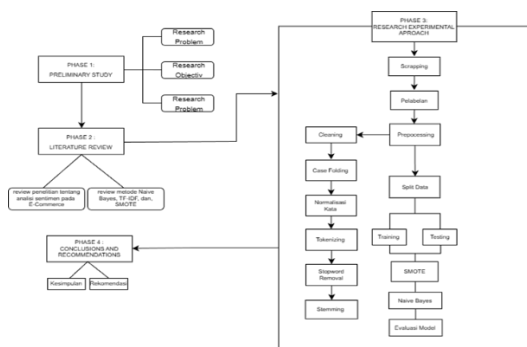
Kajian sebelumnya menunjukkan bahwa berbagai pendekatan oversampling seperti SMOTE-ENC [6], MeanRadius-SMOTE [7], serta integrasi SMOTE dengan CNN [8] telah memberikan peningkatan performa pada data tidak seimbang. Namun, sebagian besar penelitian berfokus pada dataset non-teks, membutuhkan komputasi yang tinggi, atau belum mengevaluasi integrasi teknik oversampling dengan algoritma sederhana seperti Naïve Bayes. Temuan [9] juga menegaskan bahwa tidak ada satu metode

oversampling yang unggul pada seluruh tipe data, sehingga pemilihannya harus disesuaikan dengan karakteristik dataset. Berdasarkan celah penelitian tersebut, pengembangan model analisis sentimen yang mengombinasikan algoritma Naïve Bayes dengan teknik oversampling modern seperti SMOTE pada konteks ulasan e-commerce masih diperlukan. Pendekatan ini berpotensi menghasilkan model yang lebih akurat, efisien, dan mudah diimplementasikan dalam pemrosesan ulasan pelanggan, khususnya pada platform seperti Alfagift [10].

Penelitian ini bertujuan untuk membangun model analisis sentimen yang mampu menangani ketidakseimbangan data menggunakan Naïve Bayes dan SMOTE, serta mengevaluasi peningkatannya melalui metrik akurasi, precision, recall, dan F1-score. Secara akademik, penelitian ini memperkaya literatur mengenai penanganan data tidak seimbang dan optimasi model sederhana dalam pembelajaran mesin [6], [8], [9]. Secara praktis, temuan penelitian diharapkan membantu perusahaan memahami sentimen pelanggan secara lebih akurat untuk meningkatkan strategi bisnis serta memaksimalkan kualitas layanan di era digital [11].

## MATERIALS AND METHODS

Penelitian ini merupakan studi kuantitatif dengan desain eksperimental yang menggunakan pendekatan deskriptif komparatif. Tujuan utama penelitian ini adalah menganalisis serta membandingkan performa algoritma Multinomial Naïve Bayes dalam mengklasifikasikan sentimen ulasan pengguna aplikasi Alfagift sebelum dan sesudah penerapan teknik penyeimbangan data *Synthetic Minority Oversampling Technique* (SMOTE). Pendekatan kuantitatif dipilih karena penelitian ini berfokus pada analisis berbasis numerik, pengukuran performa model secara objektif, serta evaluasi statistik melalui berbagai metrik seperti *accuracy*, *precision*, *recall*, dan *F1-score* [12], [13].



Gambar 1 Desain Penelitian

Pendekatan eksperimental diterapkan dalam penelitian ini karena analisis sentimen menuntut pengujian langsung terhadap model machine learning dengan dua konfigurasi utama, yaitu model baseline *Multinomial Naive Bayes* tanpa penyeimbangan data serta model *Multinomial Naive Bayes* yang telah dioptimalkan menggunakan *Synthetic Minority Oversampling Technique* (SMOTE). Melalui desain ini, peneliti dapat menilai secara empiris sejauh mana penerapan SMOTE berkontribusi terhadap peningkatan kemampuan model dalam mengenali kelas minoritas pada data ulasan Alfacift yang memiliki distribusi sentimen tidak seimbang [9], [14]. Evaluasi kinerja menggunakan metrik akurasi, *presisi*, *recall*, dan *F1-score* dilakukan untuk memastikan bahwa model tidak hanya menunjukkan performa tinggi pada kelas mayoritas, tetapi juga tetap responsif dan akurat dalam mendeteksi kelas minoritas, khususnya sentimen negatif.

Rancangan penelitian ini mengikuti alur eksperimen yang tersusun secara sistematis, dimulai dari tahap pengumpulan data ulasan Alfacift melalui teknik web scraping. Data yang diperoleh berupa teks tak terstruktur dalam Bahasa Indonesia, sehingga diperlukan proses pra-pemrosesan menggunakan pendekatan *Natural Language Processing* (NLP). Tahapan pra-pemrosesan tersebut meliputi *cleaning*, *case folding*, normalisasi kata, *tokenisasi*, *stopword removal*, dan *stemming*. Seluruh langkah ini berperan penting dalam menghasilkan representasi teks yang konsisten serta meminimalkan keberadaan noise pada dataset [15], [16].

Setelah teks dibersihkan, tahap berikutnya adalah transformasi data menggunakan *Term Frequency-Inverse Document Frequency* (TF-IDF) yang berfungsi mengonversi teks menjadi representasi numerik berbasis bobot kata. Representasi TF-IDF dipilih karena terbukti efektif dalam banyak penelitian analisis sentimen untuk menangkap relevansi kata dalam dokumen [17]. Selanjutnya, data

dibagi menjadi data latih dan data uji dengan proporsi 80:20. Teknik SMOTE kemudian diterapkan hanya pada data latih untuk menyeimbangkan proporsi kelas positif, negatif, dan netral tanpa menyebabkan kebocoran data ke set pengujian [18], [19].

Model Multinomial Naive Bayes kemudian dilatih menggunakan dua skenario: (1) tanpa SMOTE dan (2) dengan SMOTE. Kedua model kemudian diuji menggunakan data uji yang sama untuk memastikan perbandingan yang adil. Evaluasi dilakukan melalui confusion matrix serta metrik akurasi, presisi, recall, dan F1-score, yang direkomendasikan untuk penelitian dengan distribusi kelas tidak seimbang [12], [13]. Hasil evaluasi digunakan untuk menentukan apakah penerapan SMOTE mampu meningkatkan sensitivitas model terhadap kelas minoritas. Selain analisis kuantitatif, penelitian ini juga menyertakan interpretasi hasil klasifikasi untuk memahami pola umum opini pengguna Alfacift. Interpretasi dilakukan melalui analisis distribusi sentimen serta kata-kata dominan dalam setiap kelas untuk memberikan insight terhadap persepsi pengguna terkait layanan dan fitur Alfacift. Temuan ini memberikan kontribusi tidak hanya dalam aspek teknis pemodelan, tetapi juga dalam memahami perilaku konsumen berbasis ulasan digital.

## RESULTS AND DISCUSSION

Tahap evaluasi merupakan proses penting dalam menilai kinerja model klasifikasi yang telah dibangun, dengan tujuan untuk mengukur tingkat efektivitas model dalam menyelesaikan permasalahan analisis sentimen. Pada penelitian ini, evaluasi difokuskan untuk melihat kemampuan model dalam membedakan sentimen positif dan negatif berdasarkan data ulasan pengguna. Proses evaluasi dilakukan dengan menggunakan sejumlah metrik seperti akurasi, *precision*, *recall*, dan *F1-score*, yang masing-masing memberikan perspektif berbeda mengenai kualitas prediksi model. Melalui metrik-metrik tersebut, peneliti dapat mengetahui sejauh mana model mampu mengenali pola dalam data, menghindari kesalahan klasifikasi, serta memberikan performa yang konsisten ketika diujikan pada data baru. Dengan demikian, tahapan evaluasi tidak hanya menentukan keberhasilan model, tetapi juga menjadi dasar pertimbangan dalam melakukan perbaikan atau peningkatan performa pada proses analisis sentimen berikutnya.

### Hasil Penerapan Model Smote

Penerapan teknik SMOTE memberikan kontribusi signifikan dalam meningkatkan

kemampuan model *Naïve Bayes* untuk mengenali sentimen negatif yang sebelumnya kurang terwakili dalam dataset. Salah satu indikator keberhasilan penerapan SMOTE adalah meningkatnya nilai recall pada kelas negatif, yang menunjukkan bahwa model semakin mampu mengidentifikasi ulasan dengan sentimen negatif secara lebih akurat dan tidak mengabaikannya seperti sebelum dilakukan penyeimbangan data. Hal ini sangat penting dalam konteks analisis sentimen, terutama ketika proporsi kelas tidak seimbang. Selain akurasi, evaluasi model juga mempertimbangkan metrik seperti *precision*, *recall*, dan *F1-score* untuk memberikan gambaran yang lebih komprehensif mengenai performa model pada masing-masing kelas. Pada penelitian ini, akurasi yang diperoleh setelah penerapan model *Naïve Bayes* yang dikombinasikan dengan SMOTE mencapai 89%, menunjukkan adanya peningkatan kinerja yang cukup signifikan. Berikut ditampilkan hasil visual dari penerapan model *Naïve Bayes* + SMOTE sebagai ilustrasi evaluasi performa model.

```
Confusion Matrix:
[[393 40]
 [ 31 223]]
-----
Classification Report:
precision recall f1-score support
Negatif 0.93 0.91 0.92 433
Positif 0.85 0.88 0.86 254
accuracy 0.89 0.90 0.90 687
macro avg 0.89 0.89 0.89 687
weighted avg 0.90 0.90 0.90 687
-----
Accuracy: 0.8967
```

Gambar 2 Hasil Penerapan Model Smote

### Hasil Penerapan Model *Naïve Bayes*

Hasil ini menunjukkan bahwa tanpa bantuan model SMOTE, model *Naïve Bayes* cenderung memiliki bias ke arah kelas mayoritas, sehingga performanya kurang optimal dalam mengenali sentimen positif sebagai kelas minoritas. Ketidakseimbangan data tersebut menyebabkan nilai *recall* pada kelas positif rendah (0,77), yang berarti model sering gagal menangkap ulasan yang sebenarnya bernada positif. Dengan penerapan SMOTE pada data pelatihan, distribusi kelas menjadi lebih proporsional sehingga model dapat meningkatkan kemampuan dalam membedakan pola pada kedua kelas. Teknik oversampling ini menghasilkan sampel sintetis untuk kelas minoritas, sehingga selama proses pelatihan model tidak lagi didominasi oleh kelas mayoritas. Akibatnya, model dapat mengenali

sentimen positif dengan lebih baik dan menghasilkan nilai *recall* maupun *F1-score* yang lebih stabil. Berikut ditampilkan hasil model *Naïve Bayes* setelah proses pelatihan.

```
Confusion Matrix:
[[419 14]
 [ 59 195]]
-----
Classification Report:
precision recall f1-score support
Negatif 0.88 0.97 0.92 433
Positif 0.93 0.77 0.84 254
accuracy 0.89 0.89 0.89 687
macro avg 0.90 0.87 0.88 687
weighted avg 0.90 0.89 0.89 687
-----
Accuracy: 0.8937
```

Gambar 3 Hasil *Naïve Bayes*

### Evaluasi Kinerja Model

Evaluasi kinerja model merupakan tahap krusial untuk menilai sejauh mana algoritma *Naïve Bayes* mampu melakukan klasifikasi sentimen secara efektif. Proses evaluasi ini dilakukan dengan memanfaatkan metrik yang dihasilkan dari *Confusion Matrix*, meliputi akurasi, presisi, *recall*, dan *F1-score*. Masing-masing metrik tersebut memberikan perspektif yang berbeda mengenai kemampuan model dalam mengenali pola sentimen pada data pengujian. Untuk mendapatkan pemahaman yang lebih menyeluruh, pengujian model dilakukan pada dua kondisi, yaitu sebelum dan sesudah penerapan teknik SMOTE, sehingga pengaruh penyeimbangan data terhadap performa model dapat dianalisis secara lebih mendalam. Perbandingan kedua skenario ini memberikan landasan bagi peneliti untuk mengidentifikasi sejauh mana penyeimbangan data melalui SMOTE berkontribusi terhadap peningkatan performa klasifikasi, khususnya dalam mendeteksi kelas minoritas yang sebelumnya kurang terwakili pada dataset awal.

Hasil evaluasi awal terhadap model Multinomial *Naïve Bayes* sebelum penerapan SMOTE menunjukkan bahwa model mampu melakukan klasifikasi sentimen dengan cukup baik, ditandai dengan capaian akurasi sebesar 0.89. Pada kelas negatif, nilai precision mencapai 0.88 dan recall sebesar 0.97, menghasilkan F1-score 0.92, yang mengindikasikan bahwa model sangat efektif dalam mengidentifikasi ulasan bernuansa negatif. Namun demikian, performa pada kelas positif terlihat lebih rendah, ditandai dengan recall sebesar 0.77 meskipun precision berada pada angka 0.93, sehingga F1-score kelas positif menurun menjadi 0.84. Perbedaan performa antar kelas tersebut menunjukkan adanya kecenderungan model untuk lebih mudah mengenali sentimen negatif dibandingkan sentimen positif, yang kemungkinan

erat kaitannya dengan kondisi distribusi data yang tidak seimbang. Secara keseluruhan, nilai *macro average* dan *weighted average* berada pada kisaran 0.88-0.90, memperlihatkan bahwa meskipun model cukup stabil, sensitivitasnya terhadap masing-masing kelas belum sepenuhnya merata. Temuan ini menjadi indikasi penting bahwa teknik penyeimbangan data seperti SMOTE diperlukan untuk meningkatkan kemampuan model dalam mendeteksi kelas minoritas sehingga prediksi yang dihasilkan lebih adil, konsisten, dan representatif.

Berdasarkan hasil evaluasi setelah penerapan SMOTE, model *Multinomial Naive Bayes* menunjukkan perubahan performa yang cukup signifikan, terutama dalam kemampuan mengenali sentimen positif maupun negatif. Akurasi model meningkat menjadi 0.8967, menandakan bahwa sebagian besar prediksi sudah sesuai dengan data sebenarnya. Pada kelas negatif, *precision* naik menjadi 0.93, sedangkan *recall* mencapai 0.91, menghasilkan *F1-score* sebesar 0.92. Hal ini menunjukkan bahwa model semakin mampu mengidentifikasi ulasan negatif secara tepat dan konsisten. Sementara itu, pada kelas positif, *precision* sebesar 0.85 dan *recall* 0.88 menunjukkan keseimbangan yang baik antara ketepatan dan sensitivitas, dengan *F1-score* mencapai 0.86. Secara keseluruhan, nilai rata-rata metrik (*macro average* dan *weighted average*) sama-sama berada pada angka 0.89 hingga 0.90, yang mengindikasikan stabilitas model pada kedua kelas. Hasil ini membuktikan bahwa SMOTE tidak hanya membantu memperbaiki distribusi data, tetapi juga meningkatkan kemampuan model dalam mengenali pola sentimen, sehingga kinerjanya menjadi lebih seimbang dan tidak lagi bias terhadap kelas mayoritas.

### CONCLUSION

Berdasarkan hasil penelitian mengenai Analisis Sentimen Alfagift Menggunakan *Multinomial Naive Bayes* dan Teknik SMOTE, dapat disimpulkan beberapa hal sebagai berikut.

- 1) Berdasarkan tujuan penelitian untuk menganalisis sentimen ulasan aplikasi Alfagift, penelitian ini berhasil membangun model klasifikasi sentimen menggunakan *Multinomial Naive Bayes* yang mampu memberikan performa akurasi dan stabilitas yang baik setelah melalui tahapan pra-proses, ekstraksi fitur TF-IDF, dan evaluasi model.

- 2) Tujuan penelitian untuk mengatasi ketidakseimbangan data juga tercapai melalui penerapan SMOTE, yang terbukti meningkatkan sensitivitas model terhadap kelas minoritas serta meningkatkan nilai *recall* dan *F1-Score* tanpa menimbulkan data leakage.
- 3) Tujuan penelitian untuk menghasilkan model yang dapat mengelompokkan sentimen positif dan negatif secara efektif juga terpenuhi, ditunjukkan dengan metrik evaluasi yang konsisten dan kemampuan model dalam memproses dataset secara optimal.

### REFERENCE

- [1] O.; B. Bellar A.; Ballafkih M., "Sentiment analysis: Predicting product reviews for e-commerce recommendations using deep learning and transformers," *Mathematics*, vol. 12, no. 15, p. 2403, 2024, doi: 10.3390/math12152403.
- [2] L.; L. Yang Y.; Wang J.; Sherratt R. S., "Sentiment analysis for E-Commerce product reviews in Chinese based on sentiment lexicon and deep learning," *IEEE Access*, vol. 8, pp. 23522-23530, 2020, doi: 10.1109/ACCESS.2020.2969854.
- [3] L. Huang, "Deep learning for text sentiment analysis: A survey," *Applied and Computational Engineering*, vol. 104, pp. 135-139, 2024, doi: 10.54254/2755-2721/20241153.
- [4] C. Suhaeni and H.-S. Yong, "Mitigating class imbalance in sentiment analysis through GPT-3-generated synthetic sentences," *Applied Sciences*, vol. 13, no. 17, p. 9766, 2023, doi: 10.3390/app13179766.
- [5] T.; L. Kosolwattana C.; Hu R.; Han S.; Chen H.; Lin Y., "A self-inspected adaptive SMOTE algorithm (SASMOTE) for highly imbalanced data classification in healthcare," *BioData Min.*, vol. 16, p. Article 15, 2023, doi: 10.1186/s13040-023-00330-4.
- [6] M. Mukherjee and M. Khushi, "Smote-enc: A novel smote-based method to generate synthetic data for nominal and continuous features," *Applied System Innovation*, vol. 4, no. 1, 2021, doi: 10.3390/asi4010018.
- [7] F. Duan, S. Zhang, Y. Yan, and Z. Cai, "An Oversampling Method of Unbalanced Data for Mechanical Fault Diagnosis Based on MeanRadius-SMOTE," *Sensors*, vol. 22, no. 14, Jul. 2022, doi: 10.3390/s22145166.
- [8] J. H. Joloudari, A. Marefat, M. A. Nematollahi, S. S. Oyelere, and S. Hussain, "Effective class-

- imbalance learning based on SMOTE and convolutional neural networks," *Applied Sciences*, vol. 13, no. 6, p. 4006, 2023, doi: 10.3390/app13064006.
- [9] F. Mujahid, B. Turkoglu, E. Kaya, and T. Asuroglu, "Data oversampling and imbalanced datasets: An investigation of performance for machine learning and feature engineering," *J. Big Data*, vol. 11, p. Article 87, 2024, doi: 10.1186/s40537-024-00943-4.
- [10] B. Nemade, V. Bharadi, S. S. Alegavi, and B. Marakarkandy, "A comprehensive review: SMOTE-based oversampling methods for imbalanced classification techniques, evaluation, and result comparisons," *International Journal of Intelligent Systems and Applications in Engineering*, vol. 11, no. 9s, pp. 790–803, 2023.
- [11] V. Gooljar, T. Issa, and S. Hardin-Ramanan, "Sentiment-based predictive models for online purchases in the era of Marketing 5.0: A systematic review," *J. Big Data*, vol. 11, p. 107, 2024, doi: 10.1186/s40537-024-00947-0.
- [12] S. A. Hicks, T. Powo, and H. Booth, "On evaluation metrics for medical applications of artificial intelligence," *Sci. Rep.*, vol. 12, p. 6924, 2022, doi: 10.1038/s41598-022-09954-8.
- [13] M. Owusu-Adjei, E. Antwi, and J. Annan, "Imbalanced class distribution and performance evaluation metrics," *PLOS Digital Health*, vol. 2, no. 7, p. e0000290, 2023, doi: 10.1371/journal.pdig.0000290.
- [14] W. I. Sabilla and C. B. Vista, "Implementation of SMOTE and under sampling on imbalanced datasets for predicting company bankruptcy," *Jurnal Komputer Terapan*, vol. 7, no. 2, 2024, doi: 10.35143/jkt.v7i2.5027.
- [15] S. Gupta and G. Lehal, "A Survey on Natural Language Processing Techniques," *Information*, vol. 12, no. 9, p. 345, 2021, doi: 10.3390/info12090345.
- [16] D. W. Otter, J. R. Medina, and J. K. Kalita, "A Survey of the Usages of Deep Learning for Natural Language Processing," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 32, no. 2, pp. 604–624, 2020, doi: 10.1109/TNNLS.2020.2979670.
- [17] S. Minaee, N. Kalchbrenner, E. Cambria, N. Nikzad, M. Chenaghlu, and J. Gao, "Deep Learning-Based Text Classification: A Comprehensive Review," *ACM Comput. Surv.*, vol. 54, no. 3, pp. 1–40, 2021, doi: 10.1145/3439726.
- [18] A. H. Putra and A. Salam, "A comparative performance of SMOTE, ADASYN and random oversampling in machine-learning models on prostate cancer dataset," *Journal of Applied Informatics and Computing*, vol. 9, no. 3, 2024, doi: 10.30871/jaic.v9i3.9308.
- [19] T. Fulazzaky, A. Saefuddin, and A. M. Soleh, "Evaluating ensemble learning techniques for class imbalance in machine learning: A comparative analysis of Balanced Random Forest, SMOTE-RF, SMOTEBoost, and RUSBoost," *Scientific Journal of Informatics*, vol. 11, no. 4, 2024, doi: 10.15294/sji.v11i4.15937.