

Analisis Pengaruh Chi-Square Feature Selection terhadap Kinerja Random Forest dan XGBoost dalam Prediksi Konversi Pengunjung Website

Muhamad Rosdiana, Teti Desyani, Perani Rosyani*

Fakultas Ilmu Komputer, Program Studi Teknik Informatika, Universitas Pamulang, Kota Tangerang Selatan, Indonesia

Email: ¹dosen02354@unpam.ac.id, dosen00839@unpam.ac.id ^{2,*}dosen00837@unpam.ac.id

Email Penulis Korespondensi: dosen00837@unpam.ac.id

Submitted: 10/05/2026; Accepted: 30/06/2026; Published: 30/06/2026

Abstrak—Peningkatan jumlah pengunjung pada website *e-commerce* tidak selalu diikuti oleh peningkatan transaksi pembelian, sehingga perusahaan menghadapi kesulitan dalam mengidentifikasi pengunjung yang berpotensi melakukan konversi. Selain itu, penggunaan seluruh atribut pada dataset dapat meningkatkan kompleksitas model tanpa menjamin peningkatan kinerja prediksi. Penelitian ini bertujuan menganalisis pengaruh penerapan Chi-Square Feature Selection terhadap kinerja algoritma Random Forest dan XGBoost dalam memprediksi konversi pengunjung website. Dataset yang digunakan adalah Online Shoppers Purchasing Intention yang terdiri atas 12.330 data dengan 17 atribut prediktor dan satu atribut target. Tahapan penelitian meliputi *exploratory data analysis*, *preprocessing*, seleksi fitur menggunakan metode Chi-Square, pembangunan model klasifikasi menggunakan Random Forest dan XGBoost, serta evaluasi menggunakan Accuracy, Precision, Recall, F1-Score, Matthews Correlation Coefficient (MCC), dan Area Under Curve (AUC-ROC). Pengujian dilakukan pada empat skenario, yaitu seluruh fitur (*baseline*), Top-15, Top-10, dan Top-5 fitur hasil seleksi Chi-Square. Hasil penelitian menunjukkan bahwa penggunaan seluruh fitur menghasilkan performa terbaik dibandingkan seluruh skenario reduksi fitur. Model Random Forest *baseline* memperoleh Accuracy sebesar 90,05%, Precision 73,94%, F1-Score 63,13%, dan MCC 0,5835, sedangkan XGBoost *baseline* menghasilkan nilai AUC-ROC tertinggi sebesar 0,9271. Selain itu, atribut PageValues, BounceRates, ExitRates, ProductRelated_Duration, dan ProductRelated teridentifikasi sebagai fitur yang paling berpengaruh terhadap keberhasilan konversi pengunjung. Kontribusi penelitian ini adalah memberikan bukti empiris bahwa penerapan Chi-Square Feature Selection pada dataset Online Shoppers Purchasing Intention lebih efektif untuk mengurangi kompleksitas atribut dan mengidentifikasi fitur-fitur yang paling relevan dibandingkan meningkatkan performa klasifikasi, sehingga dapat menjadi acuan dalam pemilihan strategi seleksi fitur pada pengembangan model prediksi konversi website berbasis machine learning.

Kata Kunci: Chi-Square Feature Selection; Random Forest; XGBoost; Prediksi Konversi; Machine Learning.

Abstract—The increasing number of visitors to e-commerce websites is not always accompanied by a corresponding increase in purchase transactions, making it difficult for companies to identify visitors with high conversion potential. In addition, using all available attributes may increase model complexity without necessarily improving predictive performance. This study analyzes the impact of Chi-Square Feature Selection on the performance of Random Forest and Extreme Gradient Boosting (XGBoost) in predicting website visitor conversion. The study uses the Online Shoppers Purchasing Intention dataset consisting of 12,330 instances with 17 predictor attributes and one target attribute. The research process includes exploratory data analysis, preprocessing, Chi-Square-based feature selection, classification model development using Random Forest and XGBoost, and evaluation using Accuracy, Precision, Recall, F1-Score, Matthews Correlation Coefficient (MCC), and Area Under the Curve (AUC-ROC). Four experimental scenarios were evaluated: all features (baseline), Top-15, Top-10, and Top-5 selected features. The results show that the baseline model using all features achieved the best overall performance. The Random Forest baseline model obtained an Accuracy of 90.05%, Precision of 73.94%, F1-Score of 63.13%, and MCC of 0.5835, while the XGBoost baseline model achieved the highest AUC-ROC of 0.9271. Furthermore, PageValues, BounceRates, ExitRates, ProductRelated_Duration, and ProductRelated were identified as the most influential features affecting visitor conversion. The main contribution of this study is providing empirical evidence that Chi-Square Feature Selection is more effective in reducing feature complexity and identifying relevant attributes than improving classification performance on the Online Shoppers Purchasing Intention dataset, offering practical guidance for feature selection strategies in machine learning-based website conversion prediction.

Keywords: Chi-Square Feature Selection; Random Forest; XGBoost; Website Conversion Prediction; Machine Learning.

1. PENDAHULUAN

Perkembangan teknologi informasi telah mendorong pertumbuhan pesat sektor perdagangan elektronik (*e-commerce*) di berbagai negara [1]. Peningkatan jumlah pengunjung website tidak selalu berbanding lurus dengan peningkatan transaksi pembelian, sehingga tingkat konversi pengunjung (*conversion rate*) masih menjadi salah satu indikator utama keberhasilan sebuah platform *e-commerce* [2]. Sebagian besar pengunjung hanya melakukan aktivitas penelusuran produk tanpa menyelesaikan proses pembelian, sehingga perusahaan memerlukan strategi yang mampu mengidentifikasi calon pelanggan potensial sejak dini untuk meningkatkan efektivitas pemasaran dan personalisasi layanan [3].

Pendekatan machine learning telah banyak diterapkan untuk memprediksi kemungkinan seorang pengunjung melakukan pembelian berdasarkan perilaku selama sesi kunjungan. Kemampuan prediksi tersebut memungkinkan perusahaan memberikan rekomendasi produk, promosi, maupun penawaran khusus kepada pengguna yang memiliki peluang konversi tinggi. Oleh karena itu, pengembangan model prediksi yang akurat menjadi salah satu topik penelitian yang terus berkembang dalam bidang data mining dan kecerdasan buatan [4].

Berbagai penelitian telah dilakukan menggunakan dataset Online Shoppers Purchasing Intention [5]. Sakar dkk. menggunakan Multilayer Perceptron, Random Forest, dan Support Vector Machine dengan pendekatan oversampling dan memperoleh nilai akurasi sebesar 87,24%. Karim dan Mohsil menerapkan Random Forest dengan teknik SMOTE dan menghasilkan akurasi sebesar 86,78%. Selanjutnya, Song dan Liu menggunakan algoritma XGBoost dan memperoleh akurasi sebesar 90,15%, menunjukkan bahwa metode ensemble memiliki kemampuan yang baik dalam menangani permasalahan prediksi konversi pengunjung website. Selain itu, penelitian terbaru oleh Abdullah-All-Tanvir dkk. mengombinasikan feature selection, teknik penyeimbangan data, dan XGBoost sehingga mampu mencapai akurasi 90,65% dengan nilai Area Under ROC Curve (AUC) sebesar 0,937, yang menunjukkan performa klasifikasi yang sangat baik.

Meskipun demikian, masih terdapat beberapa kesenjangan penelitian (*research gap*) yang belum banyak dibahas pada penelitian terdahulu. Pertama, sebagian besar penelitian lebih berfokus pada peningkatan performa algoritma klasifikasi melalui teknik penyeimbangan data (*oversampling* atau SMOTE) maupun optimasi model, sementara evaluasi mengenai pengaruh metode seleksi fitur terhadap algoritma ensemble masih relatif terbatas [6]. Kedua, penelitian terdahulu umumnya hanya melaporkan peningkatan nilai akurasi tanpa melakukan analisis komprehensif mengenai perubahan performa model setelah dilakukan reduksi fitur menggunakan beberapa skenario jumlah atribut. Ketiga, belum banyak penelitian yang secara langsung membandingkan kinerja Random Forest dan XGBoost pada kondisi penggunaan seluruh fitur (*baseline*) dengan beberapa hasil seleksi fitur menggunakan metode Chi-Square (Top-15, Top-10, dan Top-5). Akibatnya, efektivitas Chi-Square Feature Selection dalam meningkatkan kualitas prediksi maupun mengurangi kompleksitas model pada dataset Online Shoppers Purchasing Intention masih belum diketahui secara jelas [7].

Berdasarkan kesenjangan tersebut, penelitian ini mengusulkan penerapan Chi-Square Feature Selection untuk menganalisis pengaruh seleksi fitur terhadap kinerja algoritma Random Forest dan XGBoost dalam memprediksi konversi pengunjung website. Evaluasi dilakukan menggunakan empat skenario, yaitu penggunaan seluruh fitur (*baseline*), Top-15, Top-10, dan Top-5 fitur hasil seleksi Chi-Square [8]. Performa model dianalisis menggunakan berbagai metrik evaluasi, meliputi Accuracy, Precision, Recall, F1-Score, Matthews Correlation Coefficient (MCC), Area Under Curve (AUC-ROC), Confusion Matrix, ROC Curve, Learning Curve, serta validasi silang menggunakan 5-Fold dan 10-Fold Cross Validation sehingga diperoleh hasil yang lebih komprehensif dibandingkan penelitian sebelumnya [9].

Kontribusi utama penelitian ini terdiri atas tiga aspek. Pertama, penelitian ini menyajikan analisis empiris mengenai pengaruh Chi-Square Feature Selection terhadap dua algoritma ensemble, yaitu Random Forest dan XGBoost, menggunakan empat skenario seleksi fitur sehingga efektivitas reduksi atribut dapat dievaluasi secara menyeluruh. Kedua, penelitian ini tidak hanya membandingkan performa model berdasarkan akurasi, tetapi juga menggunakan metrik evaluasi yang lebih lengkap seperti Precision, Recall, F1-Score, MCC, AUC-ROC, serta validasi silang untuk mengukur stabilitas model. Ketiga, penelitian ini memberikan rekomendasi mengenai strategi pemilihan fitur yang paling sesuai untuk prediksi konversi pengunjung website pada dataset Online Shoppers Purchasing Intention, sehingga hasil penelitian dapat menjadi referensi bagi pengembangan sistem prediksi konversi berbasis machine learning pada platform *e-commerce*.

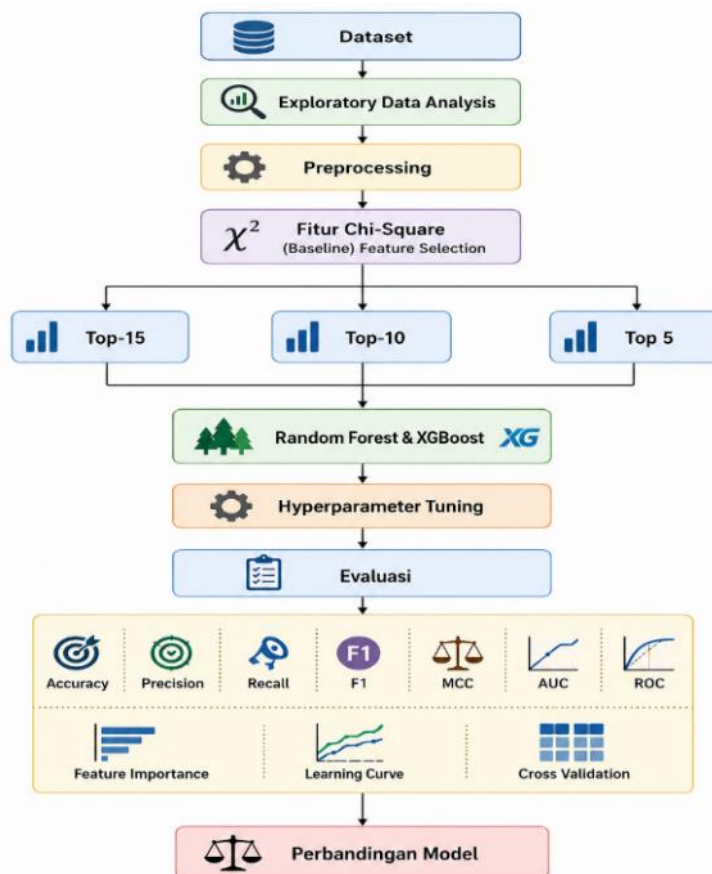
2. METODOLOGI PENELITIAN

Penelitian ini menggunakan pendekatan kuantitatif dengan metode eksperimen untuk menganalisis pengaruh penerapan Chi-Square Feature Selection [10] terhadap kinerja algoritma Random Forest [11] dan XGBoost dalam memprediksi konversi pengunjung website [12]. Dataset yang digunakan adalah Online Shoppers Purchasing Intention yang terdiri atas 12.330 data dengan 17 atribut prediktor dan satu atribut target (*Revenue*). Penelitian dilakukan melalui beberapa tahapan yang meliputi persiapan dataset, *exploratory data analysis* (EDA), *preprocessing*, seleksi fitur menggunakan metode Chi-Square, pembangunan model klasifikasi, serta evaluasi performa model. Tahapan tersebut dirancang untuk mengevaluasi pengaruh reduksi fitur terhadap kemampuan prediksi kedua algoritma ensemble secara sistematis.

Sebelum proses implementasi dilakukan, seluruh tahapan penelitian dirancang secara berurutan agar proses pengolahan data, pembangunan model, hingga evaluasi hasil dapat dilakukan secara sistematis. Tahapan penelitian yang digunakan pada penelitian ini ditunjukkan pada Gambar 1.

Gambar 1 merupakan tahapan penelitian diawali dengan proses pengumpulan dataset Online Shoppers Purchasing Intention [13], kemudian dilakukan Exploratory Data Analysis (EDA) untuk memahami karakteristik data, distribusi kelas, serta mengidentifikasi kualitas dataset sebelum dilakukan proses pemodelan. Selanjutnya dilakukan tahap *preprocessing* yang meliputi proses *label encoding* terhadap atribut kategorikal dan normalisasi data menggunakan StandardScaler sehingga seluruh atribut memiliki skala yang seragam.

Setelah proses *preprocessing*, penelitian dibagi menjadi dua skenario utama [14]. Skenario pertama menggunakan seluruh atribut sebagai baseline, sedangkan skenario kedua menerapkan Chi-Square Feature Selection untuk memperoleh atribut yang paling relevan terhadap variabel target. Berdasarkan nilai Chi-Square dilakukan pemeringkatan atribut sehingga diperoleh tiga kelompok fitur, yaitu Top-15, Top-10, dan Top-5 [15].



Gambar 1. Tahapan Penelitian

Selanjutnya setiap skenario digunakan sebagai masukan pada algoritma Random Forest dan XGBoost untuk membangun model klasifikasi. Kedua algoritma dipilih karena merupakan metode *ensemble learning* yang mampu menghasilkan performa klasifikasi yang baik pada data berdimensi tinggi serta memiliki kemampuan generalisasi yang tinggi. Untuk memperoleh konfigurasi model terbaik, dilakukan proses hyperparameter tuning sebelum tahap evaluasi.

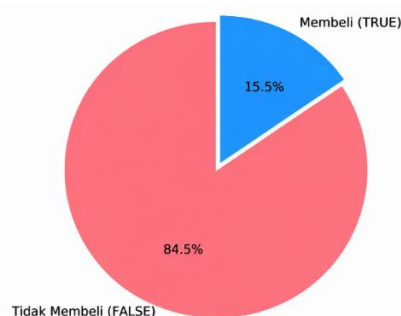
Tahap terakhir adalah evaluasi performa model menggunakan metrik Accuracy, Precision, Recall, F1-Score, Matthews Correlation Coefficient (MCC), dan Area Under Curve (AUC-ROC) [16]. Selain itu dilakukan 5-Fold dan 10-Fold Cross Validation untuk mengukur stabilitas model pada setiap skenario pengujian. Seluruh hasil evaluasi kemudian dibandingkan untuk menganalisis pengaruh penerapan Chi-Square Feature Selection terhadap performa algoritma Random Forest dan XGBoost dalam memprediksi konversi pengunjung website [17].

3. HASIL DAN PEMBAHASAN

Bagian ini menyajikan hasil implementasi metode Chi-Square Feature Selection pada algoritma Random Forest dan XGBoost dalam memprediksi konversi pengunjung website menggunakan dataset Online Shoppers Purchasing Intention. Analisis diawali dengan pembahasan karakteristik dataset dan *exploratory data analysis* (EDA) untuk memahami distribusi data serta hubungan antar atribut.[18] Selanjutnya dilakukan analisis hasil seleksi fitur menggunakan metode Chi-Square, evaluasi performa model pada berbagai skenario pengujian, serta perbandingan kinerja Random Forest dan XGBoost berdasarkan metrik Accuracy, Precision, Recall, F1-Score, Matthews Correlation Coefficient (MCC), Area Under Curve (AUC-ROC), Confusion Matrix, ROC Curve, Feature Importance, Learning Curve, dan Cross Validation. Pembahasan pada setiap tahap dilakukan untuk mengidentifikasi pengaruh seleksi fitur terhadap kemampuan model dalam memprediksi keberhasilan konversi pengunjung website secara akurat dan konsisten. [19]

3.1 Deskripsi Dataset

Tahap awal penelitian dilakukan dengan menganalisis karakteristik dataset Online Shoppers Purchasing Intention yang terdiri atas 12.330 data dengan 17 atribut prediktor dan satu atribut target yaitu Revenue. Berdasarkan hasil eksplorasi data diketahui bahwa seluruh atribut tidak memiliki nilai kosong (*missing value*) [20], sehingga proses imputasi data tidak diperlukan. Kondisi tersebut menunjukkan bahwa dataset memiliki kualitas yang baik untuk digunakan dalam proses pembangunan model klasifikasi [21].

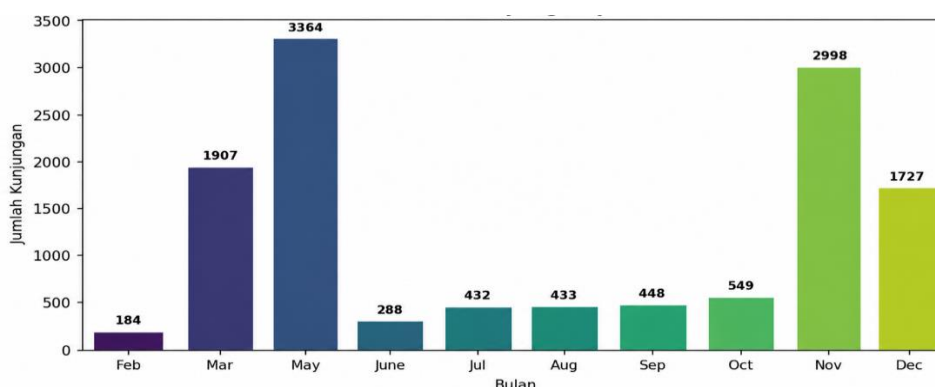


Gambar 2. Distribusi target revenue

Gambar 2 memperlihatkan distribusi variabel target yang menunjukkan bahwa jumlah data pada kelas False lebih dominan dibandingkan kelas True. Dataset tidak seimbang (imbalanced), jumlah tidak membeli: 10422 (84.5%), jumlah membeli: 1908 (15.5%). Kondisi tersebut mengindikasikan adanya ketidakseimbangan kelas (*class imbalance*) yang berpotensi memengaruhi kemampuan model dalam mengenali kelas minoritas.[22] Oleh karena itu, evaluasi model tidak hanya menggunakan metrik Accuracy, tetapi juga Precision, Recall, F1-Score, MCC, dan AUC agar performa model dapat dianalisis secara lebih komprehensif.[23]

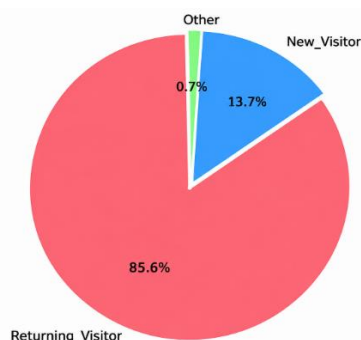
3.2 Exploratory Data Analysis (EDA)

Sebagai bagian dari tahapan Exploratory Data Analysis (EDA), dilakukan analisis terhadap distribusi data untuk mengidentifikasi pola awal yang terdapat pada dataset Online Shoppers Purchasing Intention. Analisis ini penting karena dapat memberikan gambaran mengenai perilaku pengunjung website serta membantu memahami karakteristik data sebelum dilakukan proses *preprocessing*, seleksi fitur, dan pembangunan model klasifikasi. Hasil analisis distribusi jumlah kunjungan berdasarkan bulan ditampilkan pada Gambar 3.



Gambar 3. Distribusi kunjungan perbulan

Tahap eksplorasi data dilakukan untuk memahami karakteristik setiap atribut sebelum proses pembangunan model. Gambar 3 menunjukkan distribusi jumlah sesi berdasarkan bulan kunjungan. Terlihat bahwa aktivitas pengunjung meningkat pada beberapa bulan tertentu, yang mengindikasikan adanya pengaruh musim atau periode promosi terhadap jumlah kunjungan website.



Gambar 4. Distribusi tipe pengunjung

Gambar 4 memperlihatkan bahwa tipe pengunjung Returning Visitor mendominasi dataset dibandingkan New Visitor. Hal ini menunjukkan bahwa sebagian besar pengguna merupakan pelanggan yang pernah mengakses website sebelumnya sehingga berpotensi memiliki peluang konversi yang lebih tinggi.

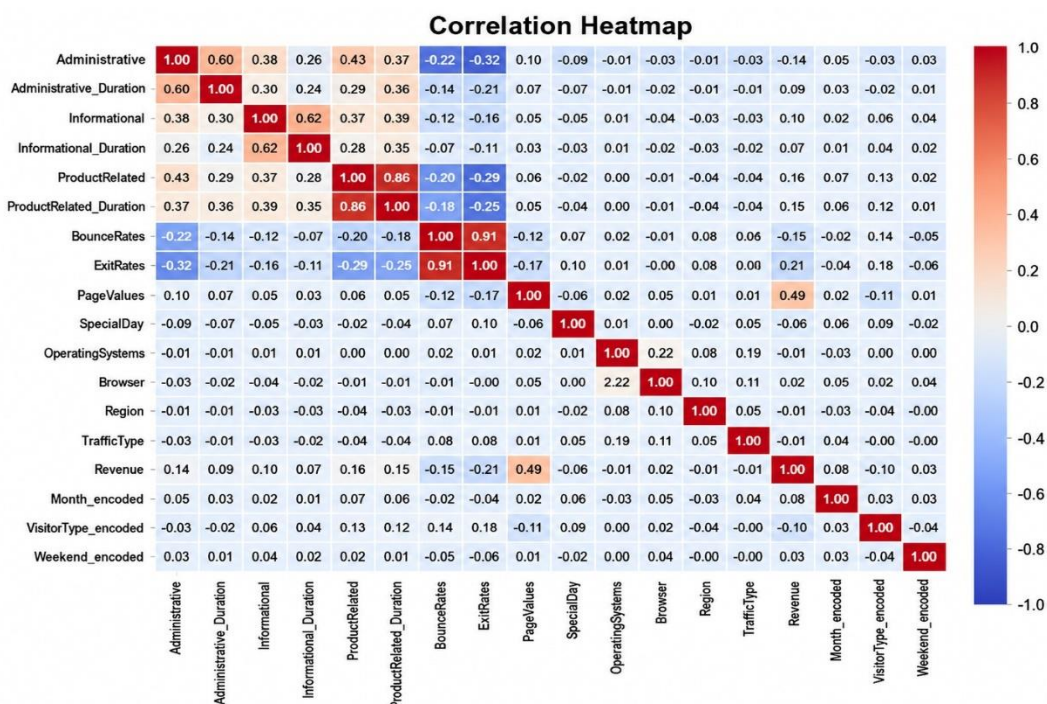
Tabel 2. Feature dan correlation

	Feature	Correlation
0	PageValues	0.492569
1	ProductRelated	0.158538
2	ProductRelated_Duration	0.152373
3	Administrative	0.138917
4	Informational	0.095200
5	Administrative_Duration	0.093587
6	Month_encoded	0.080150
7	Informational_Duration	0.070345
8	Weekend_encoded	0.029295
9	Browser	0.023984
10	TrafficType	-0.005113
11	Region	-0.011595
12	OperatingSystems	-0.014668
13	SpecialDay	-0.082305
14	VisitorType_encoded	-0.104726
15	BounceRates	-0.150673
16	ExitRates	-0.207071

Berdasarkan Tabel 2, dataset terdiri atas 12.330 data dengan 17 atribut independen dan satu atribut target (*Revenue*). Distribusi kelas menunjukkan bahwa kelas tidak membeli memiliki proporsi yang lebih besar dibandingkan kelas membeli, sehingga dataset termasuk kategori *imbalanced dataset*.

3.3 Analisis Korelasi Antar Fitur

Analisis korelasi dilakukan untuk mengetahui tingkat hubungan antara setiap atribut dengan variabel target (*Revenue*) sebagai dasar dalam memahami karakteristik data sebelum dilakukan proses seleksi fitur dan pembangunan model klasifikasi. Analisis ini menggunakan koefisien korelasi Pearson, di mana nilai positif menunjukkan hubungan searah, sedangkan nilai negatif menunjukkan hubungan berlawanan arah dengan variabel target.



Gambar 5. Heatmap Korelasi Antar Fitur

Hasil analisis korelasi terhadap variabel *Revenue* disajikan pada Tabel 2 dan divisualisasikan pada Gambar 5. Berdasarkan hasil tersebut, atribut *PageValues* memiliki nilai korelasi positif tertinggi sebesar 0,492569, yang menunjukkan bahwa semakin tinggi nilai halaman yang diperoleh selama sesi kunjungan, maka semakin besar peluang

pengunjung melakukan pembelian. Nilai korelasi tersebut juga mengindikasikan bahwa *PageValues* merupakan salah satu atribut yang paling berpengaruh dalam memprediksi konversi pengunjung *website*.

Selain *PageValues*, atribut *ProductRelated* (0,158538) dan *ProductRelated_Duration* (0,152373) juga memiliki korelasi positif terhadap variabel target. Hal ini menunjukkan bahwa semakin banyak halaman produk yang dikunjungi dan semakin lama waktu yang dihabiskan pengguna pada halaman produk, maka peluang terjadinya transaksi cenderung meningkat. Atribut *Administrative* (0,138917) dan *Informational* (0,095200) juga memiliki hubungan positif, meskipun dengan tingkat korelasi yang relatif rendah.

Sebaliknya, beberapa atribut menunjukkan korelasi negatif terhadap *Revenue*, di antaranya *ExitRates* (-0,207071), *BounceRates* (-0,150673), dan *VisitorType_encoded* (-0,104726). Nilai korelasi negatif pada *ExitRates* menunjukkan bahwa semakin tinggi persentase pengunjung yang meninggalkan *website* dari suatu halaman, maka kemungkinan terjadinya pembelian akan semakin rendah. Demikian pula dengan *BounceRates*, yang mengindikasikan bahwa pengunjung yang hanya membuka satu halaman tanpa melakukan interaksi lanjutan memiliki peluang konversi yang lebih kecil. Sementara itu, atribut seperti *TrafficType*, *Region*, dan *OperatingSystems* memiliki nilai korelasi yang sangat kecil terhadap variabel target, sehingga menunjukkan hubungan yang relatif lemah terhadap keberhasilan konversi pengunjung *website*. Kondisi ini mengindikasikan bahwa tidak seluruh atribut memberikan kontribusi yang signifikan dalam proses klasifikasi.

Secara keseluruhan, hasil analisis korelasi menunjukkan bahwa atribut yang berkaitan dengan perilaku pengguna selama mengakses *website*, seperti *PageValues*, *ProductRelated*, *ProductRelated_Duration*, *BounceRates*, dan *ExitRates*, memiliki hubungan yang lebih kuat terhadap variabel target dibandingkan atribut demografis atau teknis. Oleh karena itu, pada tahap selanjutnya dilakukan proses *Chi-Square Feature Selection* untuk mengidentifikasi atribut yang paling relevan dan mengevaluasi pengaruhnya terhadap performa algoritma *Random Forest* dan *XGBoost* dalam memprediksi keberhasilan konversi pengunjung *website*.

3.4 Hasil Chi-Square Feature Selection

Setelah melalui tahap *preprocessing*, dilakukan proses seleksi fitur menggunakan metode **Chi-Square** untuk mengidentifikasi tingkat keterkaitan setiap atribut terhadap variabel target (**Revenue**). Metode ini digunakan untuk menentukan atribut yang memiliki kontribusi paling besar dalam proses klasifikasi, di mana semakin tinggi nilai Chi-Square suatu atribut, maka semakin besar pengaruh atribut tersebut terhadap prediksi keberhasilan konversi pengunjung *website*.

Tabel 3. Ranking Nilai Chi-Square

<i>Feature</i>	<i>Chi2_Score</i>
<i>PageValues</i>	6.863.485.335
<i>BounceRates</i>	436.625.210
<i>ExitRates</i>	435.915.281
<i>ProductRelated_Duration</i>	308.960.682
<i>ProductRelated</i>	307.918.265
<i>Administrative</i>	246.262.380
<i>SpecialDay</i>	230.731.107
<i>Administrative_Duration</i>	191.448.471
<i>Informational</i>	189.577.887
<i>Informational_Duration</i>	148.512.276
<i>VisitorType_encoded</i>	35.231.548
<i>Month_encoded</i>	22.698.511
<i>Weekend_encoded</i>	15.586.645
<i>Browser</i>	2.517.836
<i>OperatingSystems</i>	2.023.218

Hasil perhitungan nilai *Chi-Square* disajikan pada Tabel 3, sedangkan visualisasi peringkat atribut ditampilkan pada Gambar 6. Berdasarkan hasil tersebut, atribut *PageValues* memperoleh nilai *Chi-Square* tertinggi sebesar 6.863.485.335, dengan selisih yang sangat signifikan dibandingkan atribut lainnya. Kondisi ini mengindikasikan bahwa *PageValues* merupakan atribut yang paling dominan dalam membedakan pengunjung yang melakukan pembelian dengan yang tidak melakukan pembelian. Semakin tinggi nilai *PageValues* yang dihasilkan selama sesi kunjungan, semakin besar peluang terjadinya konversi pada *website*.

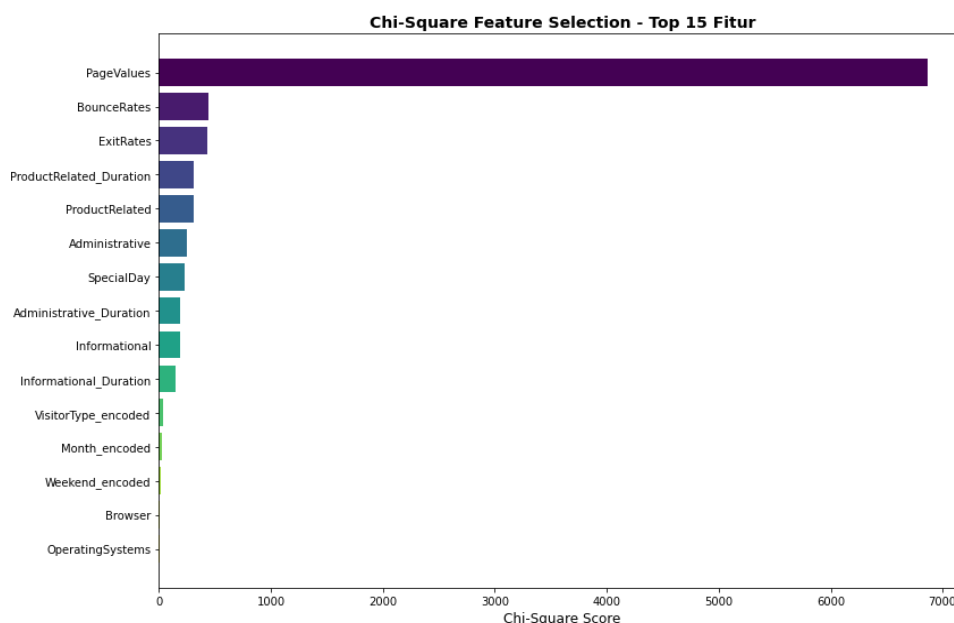
Atribut *BounceRates* dan *ExitRates* menempati peringkat berikutnya dengan nilai *Chi-Square* masing-masing sebesar 436.625.210 dan 435.915.281. Kedua atribut tersebut berkaitan dengan perilaku pengunjung selama berinteraksi dengan *website*. Nilai *BounceRates* yang tinggi menggambarkan kecenderungan pengunjung meninggalkan *website* setelah hanya membuka satu halaman, sedangkan *ExitRates* mencerminkan tingkat keluar pengguna dari halaman tertentu. Tingginya nilai *Chi-Square* pada kedua atribut tersebut mengindikasikan bahwa pola interaksi pengguna dengan *website* memiliki pengaruh yang kuat terhadap keputusan pembelian.

Selanjutnya, atribut *ProductRelated_Duration* dan *ProductRelated* juga memiliki nilai *Chi-Square* yang relatif tinggi, masing-masing sebesar 308.960.682 dan 307.918.265. Hasil ini menunjukkan bahwa durasi kunjungan pada

halaman produk serta banyaknya halaman produk yang diakses merupakan faktor penting yang dapat meningkatkan peluang konversi. Selain itu, atribut *Administrative*, *SpecialDay*, dan *Administrative_Duration* juga memberikan kontribusi terhadap proses klasifikasi, meskipun pengaruhnya tidak sebesar atribut-atribut utama.

Sebaliknya, atribut seperti *Browser*, *OperatingSystems*, *Weekend_encoded*, dan *Month_encoded* memiliki nilai *Chi-Square* yang relatif rendah. Hal tersebut menunjukkan bahwa karakteristik perangkat maupun waktu kunjungan memiliki hubungan yang lebih kecil terhadap variabel target dibandingkan atribut yang menggambarkan perilaku pengguna selama menjelajahi *website*. Dengan demikian, tidak seluruh atribut memiliki tingkat relevansi yang sama dalam proses prediksi.

Secara keseluruhan, hasil seleksi fitur menggunakan metode *Chi-Square* menunjukkan bahwa atribut yang merepresentasikan perilaku pengguna selama mengakses *website*, seperti *PageValues*, *BounceRates*, *ExitRates*, *ProductRelated*, dan *ProductRelated_Duration*, memiliki tingkat relevansi yang lebih tinggi dibandingkan atribut teknis maupun atribut demografis. Berdasarkan hasil tersebut, penelitian ini selanjutnya menggunakan tiga skenario seleksi fitur, yaitu *Top-15*, *Top-10*, dan *Top-5*, untuk menganalisis pengaruh jumlah fitur terhadap performa algoritma *Random Forest* dan *XGBoost* dalam memprediksi keberhasilan konversi pengunjung *website*.



Gambar 6. *Chi-square feature selection*

Visualisasi hasil seleksi fitur menggunakan metode *Chi-Square* ditunjukkan pada Gambar 6, yang memperlihatkan peringkat atribut berdasarkan nilai *Chi-Square* dari yang tertinggi hingga terendah. Perbedaan tersebut mengindikasikan bahwa hanya sebagian kecil fitur yang memberikan kontribusi dominan terhadap proses klasifikasi, sehingga penerapan seleksi fitur menjadi penting untuk mengurangi atribut yang kurang relevan serta meningkatkan efisiensi proses pembelajaran model.

Tabel 4. Hasil seleksi Feature

Peringkat	Top-5 Fitur	Top-10 Fitur	Top-15 Fitur
1	<i>PageValues</i>	<i>PageValues</i>	<i>PageValues</i>
2	<i>ExitRates</i>	<i>ExitRates</i>	<i>ExitRates</i>
3	<i>ProductRelated</i>	<i>ProductRelated</i>	<i>ProductRelated</i>
4	<i>BounceRates</i>	<i>BounceRates</i>	<i>BounceRates</i>
5	<i>ProductRelated_Duration</i>	<i>ProductRelated_Duration</i>	<i>ProductRelated_Duration</i>
6	-	<i>Administrative</i>	<i>Administrative</i>
7	-	<i>Administrative_Duration</i>	<i>Administrative_Duration</i>
8	-	<i>VisitorType_encoded</i>	<i>VisitorType_encoded</i>
9	-	<i>Informational</i>	<i>Informational</i>
10	-	<i>SpecialDay</i>	<i>SpecialDay</i>
11	-	-	<i>Month_encoded</i>
12	-	-	<i>Informational_Duration</i>
13	-	-	<i>Weekend_encoded</i>
14	-	-	<i>OperatingSystems</i>
15	-	-	<i>Browser</i>

Tabel 4 menunjukkan hasil seleksi fitur menggunakan metode *Chi-Square* yang digunakan dalam penelitian ini. Berdasarkan nilai *Chi-Square*, atribut *PageValues*, *ExitRates*, *ProductRelated*, *BounceRates*, dan *ProductRelated_Duration* merupakan lima fitur dengan tingkat relevansi tertinggi terhadap variabel target (*Revenue*). Selanjutnya, penelitian membentuk tiga skenario seleksi fitur, yaitu *Top-15*, *Top-10*, dan *Top-5*, untuk menganalisis pengaruh jumlah fitur terhadap performa algoritma *Random Forest* dan *XGBoost*. Skenario tersebut kemudian dibandingkan dengan penggunaan seluruh fitur (*baseline*) untuk mengetahui efektivitas metode *Chi-Square Feature Selection* dalam meningkatkan kualitas prediksi konversi pengunjung *website*..

3.5 Hasil Random Forest

Sebelum dilakukan evaluasi performa, algoritma *Random Forest* terlebih dahulu membangun sejumlah pohon keputusan berdasarkan data hasil pembagian *training set*. Setiap pohon dibangun menggunakan teknik *bootstrap sampling*, sedangkan proses pemilihan atribut pada setiap node dilakukan secara acak sehingga setiap pohon memiliki karakteristik yang berbeda. Prediksi akhir diperoleh melalui mekanisme majority voting, yaitu kelas yang paling banyak dipilih oleh seluruh pohon menjadi hasil klasifikasi akhir. Pendekatan ini mampu mengurangi variansi model dibandingkan penggunaan satu pohon keputusan sehingga menghasilkan model yang lebih stabil. Setelah proses pelatihan selesai, model diuji menggunakan empat skenario, yaitu penggunaan seluruh fitur (*baseline*), *Top-15*, *Top-10*, dan *Top-5* hasil seleksi *Chi-Square*. Sebelum penyajian hasil evaluasi, Tabel 5 memperlihatkan performa *Random Forest* pada masing-masing skenario pengujian.

Tabel 5. Hasil Random Forest

Model	Accuracy	Precision	Recall	F1-Score	MCC	AUC
RF_Semua Fitur	0.900514	0.739437	0.550699	0.631263	0.583528	0.918956
RF_Top 10	0.895107	0.707207	0.548951	0.618110	0.564451	0.897292
RF_Top 15	0.899432	0.730415	0.554196	0.630219	0.580605	0.917710
RF_Top 5	0.889430	0.681514	0.534965	0.599412	0.541644	0.890867

Berdasarkan Tabel 5, model RF_Semua Fitur menghasilkan performa terbaik dengan Accuracy 90,05%, Precision 73,94%, Recall 55,07%, F1-Score 63,13%, MCC 0,5835, dan AUC 0,9190. Sementara itu, penerapan seleksi fitur menggunakan *Top-15* masih mampu mempertahankan performa yang mendekati model dasar, sedangkan penggunaan *Top-10* dan *Top-5* menunjukkan penurunan nilai pada hampir seluruh metrik evaluasi. Hasil ini mengindikasikan bahwa semakin sedikit atribut yang digunakan, semakin besar informasi penting yang hilang sehingga kemampuan klasifikasi model ikut menurun. Secara keseluruhan, hasil tersebut menunjukkan bahwa *Random Forest* lebih optimal ketika menggunakan seluruh atribut dibandingkan hasil reduksi fitur menggunakan *Chi-Square Feature Selection*. Hal ini menunjukkan bahwa keberagaman atribut masih memberikan informasi yang relevan dalam proses pembentukan model, sehingga seleksi fitur pada dataset ini lebih berperan dalam mengurangi kompleksitas data daripada meningkatkan performa klasifikasi.

3.6 Hasil XGBoost

Evaluasi algoritma *XGBoost* dilakukan menggunakan empat skenario yang sama, yaitu seluruh fitur (*XGB_Semua Fitur*), *Top-15*, *Top-10*, dan *Top-5* hasil seleksi fitur menggunakan metode *Chi-Square*. Hasil evaluasi disajikan pada Tabel 6 berdasarkan metrik Accuracy, Precision, Recall, F1-Score, MCC, dan AUC.

Tabel 6. Hasil XGBoost

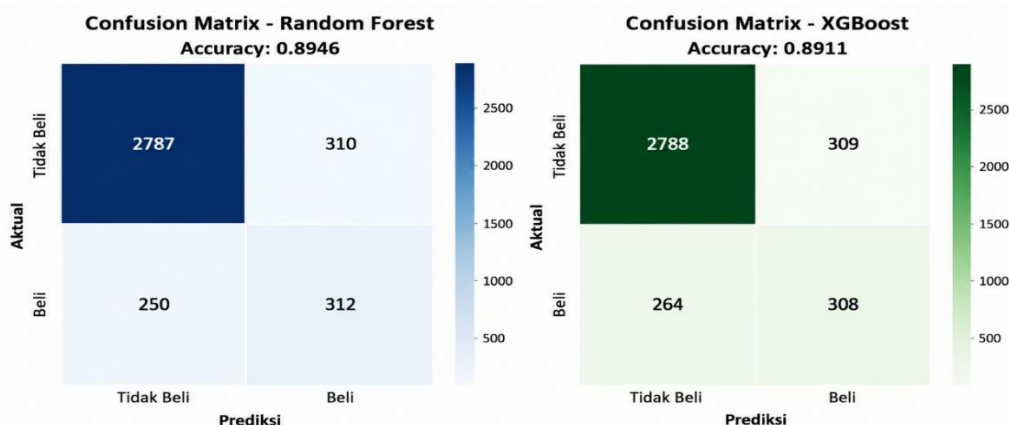
Model	Accuracy	Precision	Recall	F1-Score	MCC	AUC
XGB_Semua Fitur	0.895377	0.697228	0.571678	0.628242	0.571848	0.927062
XGB_Top 10	0.889970	0.679739	0.545455	0.605238	0.546638	0.898661
XGB_Top 15	0.894566	0.697826	0.561189	0.622093	0.566171	0.926926
XGB_Top 5	0.892403	0.697727	0.536713	0.606719	0.551924	0.892903

Berdasarkan Tabel 6, model XGB_Semua Fitur menghasilkan performa terbaik dengan Accuracy 89,54%, Precision 69,72%, Recall 57,17%, F1-Score 62,82%, MCC 0,5718, dan AUC 0,9271. Skenario *Top-15* masih menunjukkan performa yang relatif mendekati model dasar, sedangkan penggunaan *Top-10* dan *Top-5* menghasilkan penurunan pada hampir seluruh metrik evaluasi. Hasil tersebut menunjukkan bahwa penggunaan seluruh atribut masih memberikan kemampuan klasifikasi yang lebih baik dibandingkan skenario reduksi fitur. Secara keseluruhan, algoritma *XGBoost* memberikan performa terbaik ketika menggunakan seluruh atribut. Meskipun *Chi-Square Feature Selection* mampu mengidentifikasi atribut yang paling relevan, pengurangan jumlah fitur menyebabkan sebagian informasi penting tidak lagi digunakan selama proses pembelajaran model. Oleh karena itu, pada dataset ini penerapan *Chi-Square* lebih efektif untuk menyederhanakan kompleksitas data dan meningkatkan efisiensi komputasi dibandingkan meningkatkan performa klasifikasi.

3.7 Evaluasi Model

Setelah seluruh proses pelatihan model menggunakan algoritma *Random Forest* dan *XGBoost* pada setiap skenario pengujian selesai dilakukan, tahap berikutnya adalah evaluasi model untuk mengukur kemampuan klasifikasi secara

menyeluruh. Evaluasi ini bertujuan untuk mengetahui efektivitas masing-masing algoritma dalam memprediksi konversi pengunjung website, sekaligus menganalisis pengaruh penerapan Chi-Square Feature Selection terhadap performa model. Pengukuran dilakukan menggunakan berbagai metrik evaluasi, yaitu Accuracy, Precision, Recall, F1-Score, Matthews Correlation Coefficient (MCC), dan Area Under Curve (AUC-ROC). Selain itu, hasil evaluasi juga didukung oleh analisis Confusion Matrix, ROC Curve, Learning Curve, serta Cross Validation sehingga diperoleh gambaran yang lebih komprehensif mengenai tingkat akurasi, kemampuan generalisasi, dan stabilitas model pada setiap skenario pengujian.



Gambar 7. Confusion Matrix RF

Berdasarkan Gambar 7 untuk memperoleh gambaran yang lebih rinci mengenai kemampuan model dalam mengklasifikasikan setiap kelas, dilakukan analisis menggunakan **Confusion Matrix**. Analisis ini bertujuan untuk mengetahui jumlah prediksi yang benar maupun salah pada masing-masing kelas sehingga performa model tidak hanya dinilai berdasarkan nilai akurasi, tetapi juga berdasarkan kemampuan dalam mengenali kelas positif (pengunjung yang melakukan pembelian).

Berdasarkan hasil pengujian, algoritma Random Forest berhasil mengklasifikasikan sebanyak 312 data pengunjung yang benar-benar melakukan pembelian (*True Positive*), sedangkan sebanyak 260 data pembelian masih diklasifikasikan sebagai tidak melakukan pembelian (*False Negative*). Sementara itu, algoritma XGBoost menghasilkan 308 *True Positive* dan 264 *False Negative*. Hasil tersebut menunjukkan bahwa kedua algoritma memiliki kemampuan yang relatif baik dalam mengidentifikasi pengunjung yang berpotensi melakukan konversi, namun masih terdapat sejumlah data pada kelas pembelian yang belum berhasil dikenali secara optimal. Kondisi ini dipengaruhi oleh karakteristik dataset yang bersifat imbalanced, di mana jumlah data pengunjung yang tidak melakukan pembelian jauh lebih banyak dibandingkan pengunjung yang melakukan pembelian.

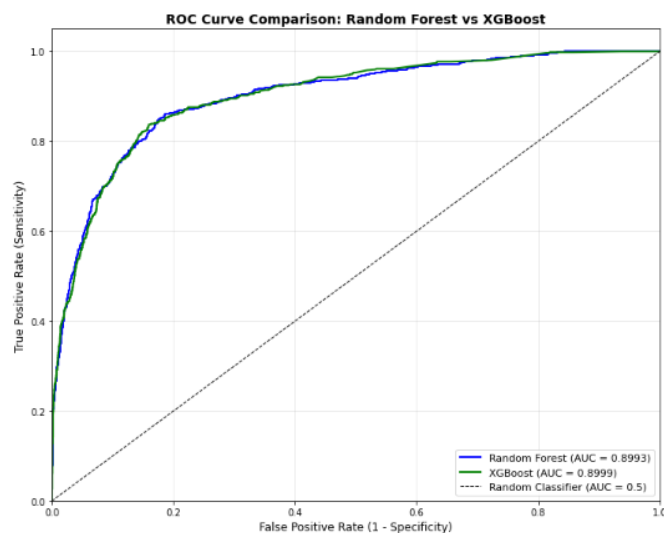
Jika dibandingkan secara langsung, Random Forest mampu menghasilkan jumlah *True Positive* yang sedikit lebih tinggi serta jumlah *False Negative* yang lebih rendah dibandingkan XGBoost. Hal ini mengindikasikan bahwa Random Forest memiliki kemampuan yang lebih baik dalam mengenali kelas positif sehingga menghasilkan nilai Recall yang sedikit lebih tinggi. Dengan demikian, Random Forest lebih efektif dalam meminimalkan kehilangan prediksi terhadap calon pelanggan yang benar-benar memiliki potensi melakukan transaksi, yang merupakan aspek penting dalam pengembangan sistem prediksi konversi pada platform *e-commerce*.

Untuk mengevaluasi kemampuan model dalam membedakan antara pengunjung yang melakukan pembelian dan yang tidak melakukan pembelian pada berbagai nilai ambang (*threshold*), dilakukan analisis menggunakan Receiver Operating Characteristic (ROC) Curve). Kurva ROC memberikan gambaran mengenai hubungan antara True Positive Rate (TPR) dan False Positive Rate (FPR), sedangkan nilai Area Under Curve (AUC) digunakan sebagai indikator kemampuan diskriminasi model. Semakin mendekati nilai 1,0 maka semakin baik kemampuan model dalam mengklasifikasikan kedua kelas. Hasil perbandingan kurva ROC untuk algoritma Random Forest dan XGBoost ditampilkan pada Gambar 8.

Berdasarkan Gambar 8, kedua algoritma menghasilkan kurva ROC yang berada jauh di atas garis diagonal, yang menunjukkan bahwa keduanya memiliki kemampuan klasifikasi yang sangat baik dalam membedakan pengunjung yang berpotensi melakukan pembelian dengan yang tidak melakukan pembelian. Nilai Area Under Curve (AUC) yang diperoleh algoritma Random Forest adalah 0,8993, sedangkan algoritma XGBoost memperoleh nilai AUC sebesar 0,8999. Nilai AUC pada kedua model termasuk dalam kategori excellent classification, karena berada di atas 0,80, sehingga menunjukkan bahwa kedua algoritma mampu menghasilkan probabilitas prediksi yang tinggi terhadap kelas target.

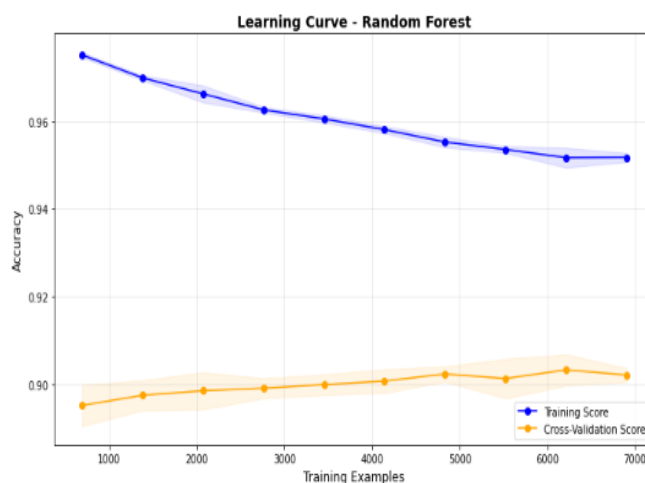
Apabila dibandingkan secara langsung, XGBoost menghasilkan nilai AUC yang sedikit lebih tinggi dibandingkan Random Forest, meskipun selisihnya relatif kecil, yaitu sebesar 0,0006. Hasil tersebut menunjukkan bahwa XGBoost memiliki kemampuan diskriminasi yang sedikit lebih baik dalam memisahkan kelas positif dan negatif pada berbagai nilai *threshold*. Keunggulan ini dipengaruhi oleh mekanisme gradient boosting yang

membangun model secara bertahap dengan memperbaiki kesalahan prediksi pada setiap iterasi, sehingga mampu menghasilkan probabilitas klasifikasi yang lebih optimal. Meskipun demikian, perbedaan nilai AUC yang sangat kecil mengindikasikan bahwa kedua algoritma memiliki tingkat performa yang hampir setara dalam memprediksi konversi pengunjung website, sehingga keduanya layak digunakan sebagai model klasifikasi pada dataset Online Shoppers Purchasing Intention.



Gambar 8. ROC Curve

Untuk mengevaluasi kemampuan generalisasi algoritma Random Forest, dilakukan analisis menggunakan Learning Curve. Analisis ini membandingkan performa model pada data pelatihan (*training score*) dan data validasi (*validation score*) sehingga dapat diketahui tingkat stabilitas model selama proses pembelajaran. Hasil *Learning Curve* algoritma Random Forest ditunjukkan pada Gambar 9.



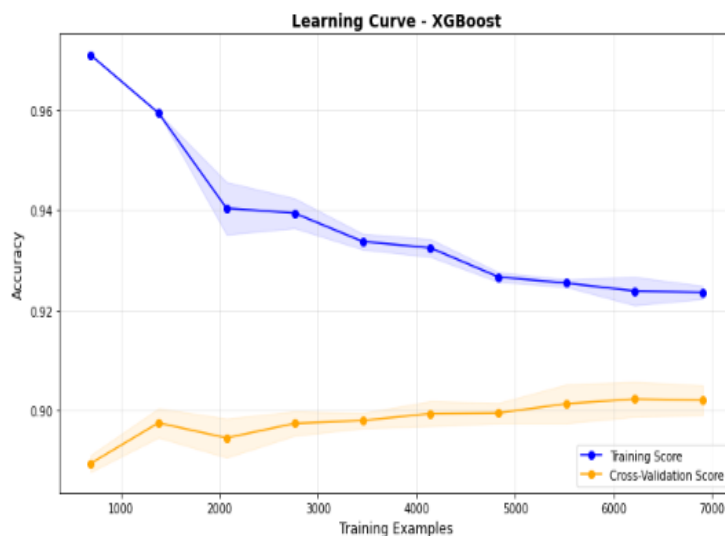
Random Forest: Model stabil ($\text{gap} = 0.0496$)

Gambar 9. Learning Curve Random Forest

Berdasarkan Gambar 9, algoritma Random Forest menunjukkan pola pembelajaran yang stabil dengan nilai *gap* antara training score dan validation score sebesar 0,0496. Selisih yang relatif kecil tersebut menunjukkan bahwa model memiliki kemampuan generalisasi yang baik dan tidak mengalami overfitting yang signifikan. Hasil ini menunjukkan bahwa Random Forest mampu mempertahankan performa secara konsisten ketika diterapkan pada data pelatihan maupun data validasi.

Selanjutnya, analisis Learning Curve juga dilakukan pada algoritma **XGBoost** untuk mengevaluasi konsistensi performa model selama proses pelatihan dan validasi. Hasil analisis ini digunakan untuk mengetahui kemampuan model dalam melakukan generalisasi terhadap data baru. Kurva pembelajaran algoritma XGBoost ditampilkan pada Gambar 10.

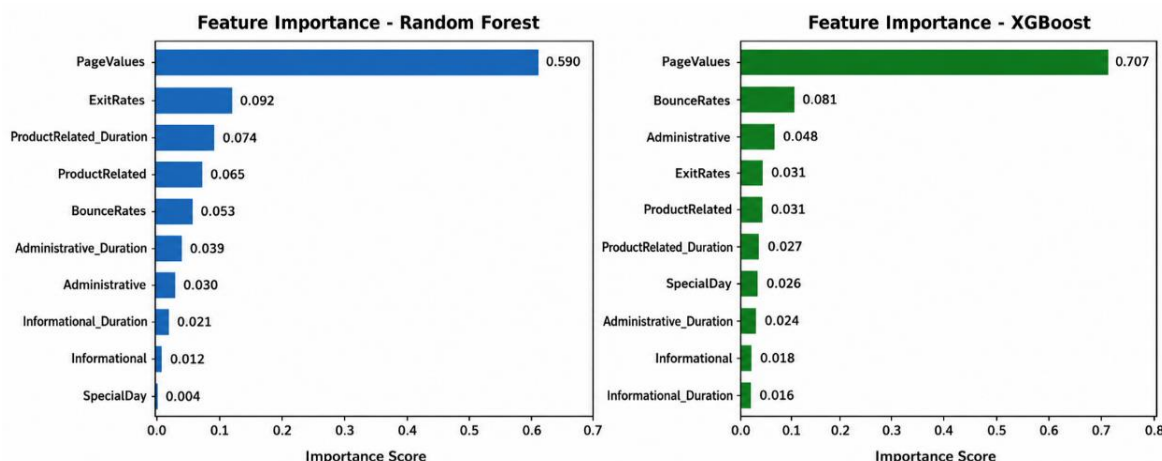
Berdasarkan Gambar 10, algoritma XGBoost menghasilkan nilai *gap* sebesar 0,0215, yang lebih kecil dibandingkan Random Forest. Nilai tersebut menunjukkan bahwa kurva training score dan validation score hampir berhimpitan sehingga model memiliki tingkat konsistensi pembelajaran yang sangat baik. Dengan demikian, XGBoost menunjukkan kemampuan generalisasi yang sedikit lebih baik dibandingkan Random Forest, meskipun kedua algoritma sama-sama menghasilkan model yang stabil untuk memprediksi konversi pengunjung website.



XGBoost: Model stabil (gap = 0.0215)

Gambar 10. Learning Curve XGBoost

Untuk mengetahui atribut yang memberikan kontribusi terbesar terhadap proses klasifikasi, dilakukan analisis Feature Importance pada algoritma Random Forest dan XGBoost. Analisis ini bertujuan untuk mengidentifikasi tingkat kepentingan setiap atribut dalam memprediksi keberhasilan konversi pengunjung website. Hasil perbandingan tingkat kepentingan atribut pada kedua algoritma ditunjukkan pada Gambar 11.



Gambar 11. Analisis Feature Importance

Berdasarkan Gambar 11, atribut PageValues menjadi fitur yang paling berpengaruh pada kedua algoritma, dengan nilai *importance* sebesar 0,589831 pada Random Forest dan 0,706615 pada XGBoost. Hasil ini menunjukkan bahwa nilai halaman yang diperoleh selama sesi kunjungan merupakan indikator utama dalam memprediksi keberhasilan konversi pengunjung website. Selain PageValues, Random Forest memberikan bobot yang lebih tinggi pada atribut ExitRates, ProductRelated_Duration, dan ProductRelated, sedangkan XGBoost lebih menekankan BounceRates, Administrative, dan ExitRates. Perbedaan tersebut menunjukkan bahwa masing-masing algoritma memiliki mekanisme pembelajaran yang berbeda dalam menentukan tingkat kepentingan setiap atribut. Secara keseluruhan, kedua algoritma secara konsisten mengidentifikasi PageValues sebagai atribut yang paling dominan dalam proses klasifikasi. Temuan ini menunjukkan bahwa perilaku pengunjung selama menjelajahi website, khususnya nilai halaman, tingkat keluar (*ExitRates*), tingkat pantulan (*BounceRates*), dan aktivitas pada halaman produk, merupakan faktor utama yang memengaruhi keberhasilan konversi pengunjung. Hasil analisis Feature Importance ini juga mendukung hasil seleksi fitur menggunakan Chi-Square, sehingga memperkuat bahwa atribut-atribut tersebut memiliki kontribusi yang signifikan dalam pembangunan model prediksi konversi berbasis *machine learning*.

3.8 Perbandingan Model

Setelah dilakukan pengujian pada seluruh skenario menggunakan algoritma *Random Forest* dan *XGBoost*, tahap selanjutnya adalah membandingkan performa masing-masing model untuk mengetahui pengaruh penerapan *Chi-Square Feature Selection* terhadap hasil klasifikasi. Perbandingan dilakukan berdasarkan metrik *Accuracy*, *Precision*,

Recall, *F1-Score*, *Matthews Correlation Coefficient (MCC)*, dan *Area Under Curve (AUC-ROC)*. Evaluasi ini bertujuan untuk mengidentifikasi model yang paling optimal dalam memprediksi keberhasilan konversi pengunjung *website* serta menganalisis efektivitas penggunaan seleksi fitur pada kedua algoritma yang digunakan.

Tabel 7 menyajikan hasil perbandingan kinerja *Random Forest* dan *XGBoost* pada empat skenario pengujian, yaitu penggunaan seluruh fitur (*baseline*), *Top-15*, *Top-10*, dan *Top-5* fitur hasil seleksi *Chi-Square*.

Tabel 7. Perbandingan semua skenario

Skenario	Model	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)	MCC	AUC-ROC
Baseline (Semua Fitur)	Random Forest	90.05	73.94	55.07	63.13	0.5835	0.9190
Baseline (Semua Fitur)	XGBoost	89.54	69.72	57.17	62.82	0.5718	0.9271
Top-15 Fitur Chi-Square	Random Forest	89.94	73.04	55.42	63.02	0.5806	0.9177
Top-15 Fitur Chi-Square	XGBoost	89.46	69.78	56.12	62.21	0.5662	0.9269
Top-10 Fitur Chi-Square	Random Forest	89.51	70.72	54.90	61.81	0.5645	0.8973
Top-10 Fitur Chi-Square	XGBoost	89.00	67.97	54.55	60.52	0.5466	0.8987
Top-5 Fitur Chi-Square	Random Forest	88.94	68.15	53.50	59.94	0.5416	0.8909
Top-5 Fitur Chi-Square	XGBoost	89.24	69.77	53.67	60.67	0.5519	0.8929

Berdasarkan tabel 7 seluruh hasil pengujian, dapat disimpulkan bahwa penggunaan seluruh fitur (*baseline*) masih menjadi pendekatan terbaik untuk dataset *Online Shoppers Purchasing Intention*. Penerapan *Chi-Square Feature Selection* pada skenario *Top-15*, *Top-10*, dan *Top-5* belum mampu meningkatkan performa klasifikasi baik pada algoritma *Random Forest* maupun *XGBoost*. Oleh karena itu, model *Random Forest baseline* ditetapkan sebagai model terbaik dalam penelitian ini dengan nilai *Accuracy* sebesar 90,05%, sedangkan *XGBoost baseline* menghasilkan kemampuan diskriminasi terbaik dengan nilai *AUC-ROC* sebesar 0,9271

Hasil penelitian menunjukkan bahwa penerapan *Chi-Square Feature Selection* tidak selalu meningkatkan performa model klasifikasi. Pada dataset *Online Shoppers Purchasing Intention*, atribut yang memiliki nilai *Chi-Square* rendah masih memberikan kontribusi terhadap proses prediksi sehingga pengurangan jumlah fitur justru menyebabkan penurunan performa. Temuan ini menunjukkan bahwa efektivitas seleksi fitur sangat bergantung pada karakteristik dataset dan algoritma yang digunakan.

3.9 Pembahasan

Hasil penelitian menunjukkan bahwa penggunaan seluruh fitur (*baseline*) menghasilkan performa yang lebih baik dibandingkan seluruh skenario reduksi fitur menggunakan *Chi-Square Feature Selection*. *Random Forest* memperoleh nilai *Accuracy* sebesar 90,05%, sedangkan *XGBoost* menghasilkan nilai *AUC* sebesar 0,9271, yang menunjukkan kemampuan diskriminasi model yang sangat baik. Temuan ini mengindikasikan bahwa pada **dataset** *Online Shoppers Purchasing Intention*, sebagian besar atribut masih memberikan informasi yang relevan terhadap proses klasifikasi sehingga pengurangan atribut belum mampu meningkatkan performa model.

Hasil penelitian ini memiliki kesamaan dengan penelitian mehdi et al [11] yang melaporkan bahwa algoritma *XGBoost* mampu menghasilkan akurasi di atas 90% pada prediksi konversi pengunjung *website*. Kesamaan tersebut menunjukkan bahwa metode *ensemble boosting* memiliki kemampuan yang baik dalam mempelajari pola perilaku pengunjung *website* yang kompleks. Namun demikian, penelitian ini memberikan tambahan informasi bahwa penerapan *Chi-Square Feature Selection* tidak selalu meningkatkan performa model meskipun mampu mengurangi jumlah atribut.

Penelitian ini juga sejalan dengan penelitian Abdullah-All-Tanvir et al. [5] yang menyatakan bahwa seleksi fitur dapat meningkatkan efisiensi proses pembelajaran model. Akan tetapi, penelitian tersebut mengombinasikan *feature selection* dengan teknik penyeimbangan data sehingga memperoleh peningkatan performa yang lebih tinggi. Sebaliknya, penelitian ini secara khusus mengevaluasi pengaruh *Chi-Square* tanpa melakukan modifikasi distribusi kelas sehingga pengaruh reduksi fitur dapat diamati secara lebih objektif.

Jika dibandingkan dengan penelitian Sakar et al., [24] *Random Forest* pada penelitian ini menghasilkan nilai *Accuracy* yang lebih tinggi. Hal tersebut diduga dipengaruhi oleh proses *hyperparameter tuning* dan evaluasi menggunakan beberapa skenario seleksi fitur sehingga konfigurasi model menjadi lebih optimal. Selain itu, penelitian Ichsan dkk [25] menggunakan pendekatan *SMOTE* untuk menangani ketidakseimbangan kelas, sedangkan penelitian ini tidak menerapkan teknik *oversampling*. Walaupun demikian, performa yang diperoleh tetap kompetitif sehingga menunjukkan bahwa *Random Forest* masih memiliki kemampuan generalisasi yang baik pada dataset asli.

Secara keseluruhan, penelitian ini memberikan kontribusi berupa analisis empiris mengenai efektivitas *Chi-Square Feature Selection* terhadap dua algoritma *ensemble* melalui empat skenario seleksi fitur yang berbeda. Hasil penelitian membuktikan bahwa reduksi fitur lebih bermanfaat untuk mengurangi kompleksitas model dibandingkan meningkatkan performa klasifikasi. Selain itu, atribut *PageValues*, *BounceRates*, *ExitRates*, *ProductRelated*, dan *ProductRelated_Duration* terbukti merupakan faktor yang paling berpengaruh dalam menentukan keberhasilan konversi pengunjung *website* sehingga dapat menjadi acuan dalam pengembangan sistem prediksi konversi pada platform *e-commerce*.

4. KESIMPULAN

Penelitian ini berhasil menganalisis pengaruh penerapan Chi-Square Feature Selection terhadap kinerja algoritma Random Forest dan XGBoost dalam memprediksi konversi pengunjung website menggunakan dataset Online Shoppers Purchasing Intention. Hasil penelitian menunjukkan bahwa penerapan seleksi fitur mampu mengidentifikasi atribut-atribut yang paling relevan terhadap variabel target serta mengurangi kompleksitas data, meskipun tidak memberikan peningkatan performa klasifikasi dibandingkan penggunaan seluruh atribut. Berdasarkan hasil evaluasi, kedua algoritma mampu menghasilkan model dengan tingkat akurasi, kemampuan diskriminasi, dan stabilitas yang baik, sehingga layak digunakan untuk mendukung prediksi konversi pengunjung website. Selain itu, atribut PageValues, ExitRates, BounceRates, ProductRelated, dan ProductRelated_Duration terbukti menjadi faktor yang paling berpengaruh dalam menentukan keberhasilan konversi, sehingga dapat dijadikan dasar dalam pengembangan strategi personalisasi layanan dan pemasaran digital pada platform *e-commerce*. Kontribusi utama penelitian ini adalah memberikan bukti empiris mengenai efektivitas penerapan Chi-Square Feature Selection pada dua algoritma *ensemble learning*, yaitu Random Forest dan XGBoost, melalui beberapa skenario seleksi fitur sehingga diperoleh pemahaman yang lebih komprehensif mengenai hubungan antara reduksi atribut, kompleksitas model, dan performa klasifikasi. Hasil penelitian ini juga memberikan rekomendasi bahwa penggunaan seluruh atribut masih menjadi pilihan terbaik pada dataset yang digunakan, sedangkan penerapan Chi-Square lebih tepat dimanfaatkan untuk menyederhanakan model tanpa mengurangi kemampuan prediksi secara signifikan. Temuan tersebut diharapkan dapat menjadi referensi bagi peneliti maupun praktisi dalam mengembangkan sistem prediksi konversi pengunjung website yang lebih efisien, akurat, dan mudah diimplementasikan. Penelitian selanjutnya disarankan mengombinasikan metode seleksi fitur dengan teknik penanganan *imbalanced dataset* atau pendekatan optimasi lainnya untuk memperoleh performa model yang lebih optimal.

REFERENCES

- [1] R. Kassemeyer, S. Alavi, and J. Habel, "Implementing e-commerce channels in business-to-business selling," *International Journal of Research in Marketing*, 2025, doi: 10.1016/j.ijresmar.2025.10.003.
- [2] R. G. Mahendra, R. A. Trenady, and P. T. Pungkasanti, "Penerapan Metode Analytical Hierarchy Process dan Additive Ratio Assessment Dalam Menentukan Target Promosi Universitas," *Building of Informatics, Technology and Science (BITS)*, vol. 6, no. 2, Sep. 2024, doi: 10.47065/bits.v6i2.5469.
- [3] P. Jansen, M. Colley, T. Pfeifer, and E. Rukzio, "Visualizing imperfect situation detection and prediction in automated vehicles: Understanding users' perceptions via user-chosen scenarios," *Transp. Res. Part F Traffic Psychol. Behav.*, vol. 104, pp. 88–108, Jul. 2024, doi: 10.1016/j.trf.2024.05.015.
- [4] K. Mili, I. Bengana, and M. S. Benmoussa, "Integrating intercultural communication into E-commerce theory: The DCEM framework for online shopping behavior," Dec. 01, 2025, *Elsevier B.V.* doi: 10.1016/j.chbr.2025.100810.
- [5] Abdullah-All-Tanvir, I. Ali Khandokar, A. K. M. Muzahidul Islam, S. Islam, and S. Shatabda, "A gradient boosting classifier for purchase intention prediction of online shoppers," *Heliyon*, vol. 9, no. 4, Apr. 2023, doi: 10.1016/j.heliyon.2023.e15163.
- [6] C. Karunakaran, V. Niranjana, and A. S. Setlur, "Random Forest and XGBoost-based ensemble models for colorectal cancer exome variant classification and web application deployment for early prediction," *Computational and Structural Biotechnology Reports*, vol. 2, p. 100063, Jan. 2025, doi: 10.1016/j.csbr.2025.100063.
- [7] I. M. Rajagukguk, R. Hartanto, Julian, and R. Halim, "Comparative Analysis of XGBoost, Random Forest, and Logistic Regression for Classifying Jakarta's Air Pollution Index (ISPU)," in *Procedia Computer Science*, Elsevier B.V., 2025, pp. 108–120. doi: 10.1016/j.procs.2025.08.264.
- [8] R. K. E. Tau, A. Yahya, M. Mangwala, and N. M. J. Ditshego, "XGBoost–random forest stacking with dual-state Kalman filtering for real-time battery SOC estimation," *Results in Engineering*, vol. 27, Sep. 2025, doi: 10.1016/j.rineng.2025.106428.
- [9] F. Zare Ebrahimbabad, H. Yazdani, A. Hakim, and M. Asarian, "Augmented Reality Versus Web-Based Shopping: How Does AR Improve User Experience and Online Purchase Intention," *Telematics and Informatics Reports*, vol. 15, Sep. 2024, doi: 10.1016/j.teler.2024.100152.
- [10] M. Fikri, R. Herteno, R. A. Nugroho, S. W. Saputro, and F. Abadi, "Multi-Criteria Decision Making Dalam Seleksi Fitur Ensemble Untuk Prediksi Cacat Perangkat Lunak Multi-Criteria Decision Making In Ensemble Feature Selection," vol. 12, no. 6, pp. 1385–1394, 2025.
- [11] S. T. Hamidou and A. Mehdi, "Enhancing IDS performance through a comparative analysis of Random Forest, XGBoost, and Deep Neural Networks," *Machine Learning with Applications*, vol. 22, p. 100738, Dec. 2025, doi: 10.1016/j.mlwa.2025.100738.
- [12] S. S. Shafin, "An explainable feature selection framework for web phishing detection with machine learning," *Data Science and Management*, vol. 8, no. 2, pp. 127–136, Jun. 2025, doi: 10.1016/j.dsm.2024.08.004.
- [13] X. Wei, "Research on Preprocessing Techniques for Software Defect Prediction Dataset Based on Hybrid Category Balance and Synthetic Sampling Algorithm," *Procedia Comput. Sci.*, vol. 262, pp. 840–848, 2025, doi: 10.1016/j.procs.2025.05.117.
- [14] D. P. Sakas, D. P. Reklitis, N. T. Giannakopoulos, and P. Trivellas, "The influence of websites user engagement on the development of digital competitive advantage and digital brand name in logistics startups," *European Research on Management and Business Economics*, vol. 29, no. 2, May 2023, doi: 10.1016/j.iedeen.2023.100221.
- [15] L. W. Hiselius, E. Adell, U. Berggren, and L. S. Rosqvist, "Consumer involvement in last-mile e-commerce transport and impact on travel Patterns: A case study in Sweden," *Transp. Res. Part A Policy Pract.*, vol. 211, Sep. 2026, doi: 10.1016/j.tra.2026.105087.
- [16] W. A. Srisathan, C. Ketkaew, N. Jantuma, and P. Naruetharadhol, "Trust and website conversion in consumer responses to green product purchasing: A new perspective through the lens of innovative website Design's technology integration," *Heliyon*, vol. 10, no. 1, Jan. 2024, doi: 10.1016/j.heliyon.2023.e23764.



- [17] H. Khani Sanij, R. Babagoli, and R. Mohammadi Elyasi, “Enhancing stone matrix asphalt performance with sugarcane bagasse ash: Mechanical properties and machine learning-based predictions using XGBoost and random forest,” *Case Studies in Construction Materials*, vol. 23, Dec. 2025, doi: 10.1016/j.cscm.2025.e05186.
- [18] M. He, T. D. Pham Thi, H. M. Tran, and N. T. Duong, “Consumers’ intentions to use online shopping apps: A comparative analysis,” *Acta Psychol. (Amst)*, vol. 259, Sep. 2025, doi: 10.1016/j.actpsy.2025.105414.
- [19] J. Q. Ruan, Z. S. Chen, Y. Yang, and M. Deveci, “How the metaverse enhances e-commerce supply chain resilience: An actor-network theory perspective,” *Advanced Engineering Informatics*, vol. 74, Sep. 2026, doi: 10.1016/j.aei.2026.104710.
- [20] N. J. Downing, “Missing value imputation in environmental, social, and governance data: an impact on emissions scores,” *Financ. Res. Lett.*, vol. 85, pp. 1–10, Nov. 2025, doi: 10.1016/j.frl.2025.107818.
- [21] M. Zhaifir and N. Rukhiviyanti, “Analisis Perbandingan Algoritma Machine Learning untuk Prediksi Konversi Penjualan UMKM Shiroapparel pada Platform Social Commerce Di Tik Tok Shop A Comparative Study of Machine Learning Algorithms for Predicting Sales Conversion of Shiroapparel MSMEs on the TikTok Shop Social Commerce Platform.”
- [22] A. Shoukat, Z. Liu, Y. Y. A. Abuker, J. Li, and L. Mao, “Dual attention-enhanced data augmentation for diagnosing water management faults in proton exchange membrane fuel cells using imbalanced multi-sine AC data,” *Energy and AI*, vol. 22, Dec. 2025, doi: 10.1016/j.egyai.2025.100630.
- [23] J. R. Pivin-Bachler and E. L. van den Broek, “SIMBA: A robust and generalizable measure of data imbalance,” *Patterns*, vol. 6, no. 12, Dec. 2025, doi: 10.1016/j.patter.2025.101395.
- [24] A. Balazs, S. Tuominen, and A. Kangas, “Enhancing forest inventory Accuracy: Comparing 3D-CNN and k-NN with genetic algorithm Approaches using ALS data across boreal bioregions,” *Comput. Electron. Agric.*, vol. 237, Oct. 2025, doi: 10.1016/j.compag.2025.110576.
- [25] N. Ichsan, R. Sopandi, H. Priyandaru, and M. Tabrani, “Pendekatan Level Data Smote Pada Algoritma Bagging C4.5 Untuk Prediksi Cacat Software Smote Data Level Approach of C4.5 Bagging Algorithm for Software Defect Prediction,” *Cermin: Jurnal Penelitian*, vol. 7, pp. 402–416, 2023.