ISSN 2723-6471

Enhanced Detection of Consumer Behavioral Shifts in E-Commerce Platforms with Transformer-Based Algorithms

Rahmad B.Y Syah[1], Marischa Elveny[2,*], Soleh Darmansyah[3], Lia Silviana[4]

[1,3]Excellent Centre of Innovations and New Science-PUIN, Faculty of Engineering, Universitas Medan Area, Medan, 20223. Indonesia

[2,4]Fasilkom-TI, Universitas Sumatera Utara, Medan, 20115. Indonesia

(Received: March 22, 2025; Revised: June 05, 2025; Accepted: August 19, 2025; Available online: September 11, 2025)

Abstract

This research aims to analyze changes in consumer behavior on e-commerce platforms using consumer interaction data such as view, add to cart, and purchase. Identifying changes in consumer behavior on e-commerce platforms is very important because it can provide deeper insight into consumer motivations and preferences. By better understanding how consumers interact with products, companies can design more targeted strategies to increase conversions, reduce cart abandonment, and improve the overall customer experience. The DistilBERT based prediction model is applied to detect and predict changing patterns of consumer behavior in the purchasing process. DistilBERT was chosen because of its more efficient capabilities compared to previous models which enable faster data processing and lower resource usage, which is very important for real-time applications on e-commerce platforms with big data. The data used includes consumer interactions during a certain period, with model evaluation using precision, recall, F1-score, and accuracy metrics. The results showed that despite an increase in the number of actions such as View and Add to Cart, conversion to Purchase was still hampered, indicating a cart abandonment problem. The model used managed to achieve 90% accuracy, with a precision value of 0.87, recall of 0.85, and F1-score of 0.86, showing excellent performance in predicting changes in consumer behavior. Based on the results of this analysis, companies can optimize marketing strategies by targeting consumers who have added products to their basket but have not yet made a purchase, as well as making price adjustments, discounts, and limited time offers. This research also emphasizes the importance of using real-time data to dynamically adjust marketing strategies and improve customer experience.

Keywords: Detection, Consumer Behavior, E-Commerce Platforms, Transformer-Based Algorithms, DistilBERT, Prediction

1. Introduction

E-commerce has emerged as a significant industry in the digital economy, experiencing fast expansion in recent years [1], [2], [3]. One of the most difficult difficulties for e-commerce systems is comprehending constantly changing consumer behavior. Changes in consumer behavior, such as purchase patterns, product preferences, and platform engagement levels, can have an impact on marketing campaigns, product suggestions, and inventory management [4], [5]. E-commerce platforms frequently collect a lot of information on user behavior, such as clicks, product searches, cart additions, transactions, and product reviews. Despite this amount of data, it is challenging to discover and predict real-time behavioral shifts. Identifying these behavioral adjustments is crucial for increasing personalization since it allows the platform to deliver more precise recommendations and improve the user experience [6], [7]. E-commerce platforms can optimize marketing tactics by offering appropriate offers or reminders to consumers who have been identified as undergoing a shift in behavior [8], [9].

As deep learning progresses, transformer-based algorithms like DistilBERT do well in processing sequential and contextual data, particularly for Natural Language Processing (NLP) tasks [10], [11], [12]. In the area of e-commerce, the DistilBERT model outperforms standard models in terms of learning consumer behavior patterns and feature associations. then used the confusion matrix to assess the effectiveness of the classification model [13], [14], [15]. Previous research [16] concentrated on the use of conventional classification algorithms like Random Forest and Support Vector Machines (SVM) to forecast customer purchasing behavior based on e-commerce interaction data. The

***Corresponding author: Marischa Elveny (marischaelveny@usu.ac.id)**
DOI: https://doi.org/10.47738/jads.v6i4.907

ISSN 2723-6471

model takes advantage of features like transaction history, customer demographics, and products viewed. These models are typically restricted to current features and cannot reflect more complicated sequential or temporal situations. In addition, [17] This study employs Bidirectional Encoder Representations from Transformers (BERT) to anticipate consumer behavior on e-commerce platforms. BERT is used to interpret customer action sequences in a deeper context. BERT requires a lot of resources for training and is often slower and more expensive in real-world applications than lighter models like DistilBERT [18]. Although several lightweight models such as TinyBERT and ALBERT were considered, preliminary benchmarking revealed that DistilBERT strikes an optimal balance between inference speed and classification accuracy [19]. Compared to TinyBERT, DistilBERT achieved slightly higher F1-scores (+2%) with minimal increase in latency. Meanwhile, ALBERT's reduced parameter sharing introduced slight underfitting in shorter sequences. Thus, DistilBERT was chosen based on empirical performance and its robust documentation for e-commerce NLP tasks [20].

2. Literature Review

In this study, a behavioral shift is defined as a change in the typical sequence or frequency of consumer actions over time particularly deviations from the common pattern of View → Add to Cart → Purchase. These shifts may manifest as increased viewing with no subsequent cart action, frequent cart additions without purchases (indicating cart abandonment), or abrupt skipping of intermediate actions (e.g., direct purchase after view) [21]. Furthermore, we categorize behavioral shifts by temporal stage in the customer lifecycle (early interest vs purchase intent) and by action sequence divergence from historical patterns [22].

The author's contribution based on previous research, models such as Random Forest and SVM have been used to detect changes in consumer behavior but have limitations in real-time behavior detection, especially on big data that is often found on e-commerce platforms [23]. These limitations are mainly related to the ability of these models to handle sequential and temporal data, as well as to process large amounts of text data. DistilBERT was chosen in this study because of its ability to be more efficient in processing big data faster and requiring fewer resources compared to previous models such as BERT. DistilBERT is more suitable for detecting changes in consumer behavior in real-time on e-commerce platforms, which require fast processing to optimize user experience and marketing strategies [24]. Research gap and Contribution Consumer Behavioral Shifts in E-Commerce Platforms can be seen in table 1.

Table 1. Research gap and Contribution Consumer Behavioral Shifts in E-Commerce Platforms

Source	Research Gap	Contribution of Method
Karabila et al. [17]	Employing BERT for behavioural prediction yields favourable contextual outcomes; yet it is excessively resource-intensive for large-scale e-commerce systems.	BERT with DistilBERT (~40% smaller), maintains lower latency accuracy.
Zhang & Shafiq [10]	Transformer reviews have not specifically applied changes in consumer behavior in e-commerce.	DistilBERT to detect shifts in specific e-commerce consumer behaviour (View → Add to Cart → Purchase).
Beránek & Merunka [15]	Focus on comparing large NLP models in BI	DistilBERT-based NLP integration with SHAP interpretation in e-commerce domain prediction decisions.
Alantari et al. [20]	Comparison of transformer models does not consider efficiency vs accuracy for deployment.	DistilBERT achieves the best balance compared to TinyBERT and ALBERT for consumer behaviour prediction.

3. Methodology

The research methodology includes data collected from customer interactions on e-commerce platforms [25]. This includes user actions such as viewing products (View), adding products to cart (Add to Cart), and making purchases (Purchase). Although the system is currently tested using batch historical data, the model is designed with real-time deployment in mind. To support real-time inference, the model is integrated into a lightweight inference pipeline with an average prediction latency of < 250 ms per user sequence using GPU-based serving. For future deployment, we plan to integrate this pipeline with streaming platforms (e.g., Apache Kafka or AWS Kinesis) to facilitate real-time data ingestion, processing, and behavioral shift detection. Next, prepare the data for the model, including cleaning, feature transformation, and tokenization.

Develop a model to detect and predict changes in consumer behavior using DistilBERT. Tokenization is applied to text data, such as product reviews, and categorical data, such as product categories, to ensure that the data is acceptable to the DistilBERT model. Tokenization is done by dividing the text into smaller units (tokens), which are then processed to produce a numeric representation that matches the input required by DistilBERT. Additionally, temporal features (such as timestamps) are processed to capture dynamic consumer behavior patterns, such as changes in preferences and interactions that are influenced by time. To handle this time information, a time normalization technique is used to ensure that the temporal data can be analyzed properly, identifying short-term and long-term trends in consumer behavior. This process allows the model to predict behavioral changes more accurately, accounting for fluctuations in consumer behavior that occur over time. To handle heterogeneous inputs, we processed textual reviews through the standard DistilBERT tokenizer pipeline. For categorical actions (View, Add to Cart, Purchase) and product metadata (e.g., category), we used embedding vectors that were concatenated with the CLS token output of DistilBERT. This multimodal fusion allowed the model to incorporate both natural language context and structured behavior signals. Categorical features were encoded using one-hot or label encoding depending on cardinality.

The selection of evaluation metrics in this study, including accuracy, precision, recall, and F1-score, was chosen to ensure a comprehensive evaluation of the model's performance in detecting changes in consumer behavior. To improve the interpretability of the model, SHAP (SHapley Additive Explanations) was used to explain the contribution of each feature in the model's decision-making process, applied to sequential and temporal data to provide a deeper understanding of the influence of variables such as interaction time and number of clicks on consumer behavior decisions. The use of SHAP in the context of temporal data allows for the identification of the most influential features in predicting whether consumers will make a purchase after viewing or adding a product to the cart. Categorical features such as Product Category were encoded using one-hot encoding before being embedded into the numerical input space. The one-hot vectors were passed through a dense transformation layer, allowing them to be merged with DistilBERT's CLS token output for joint SHAP interpretation. This ensured that the feature attribution remains interpretable even in the presence of discrete variables. As well as assessing the model's performance to predict shifts in consumer behavior using relevant metrics. can be seen in figure 1.

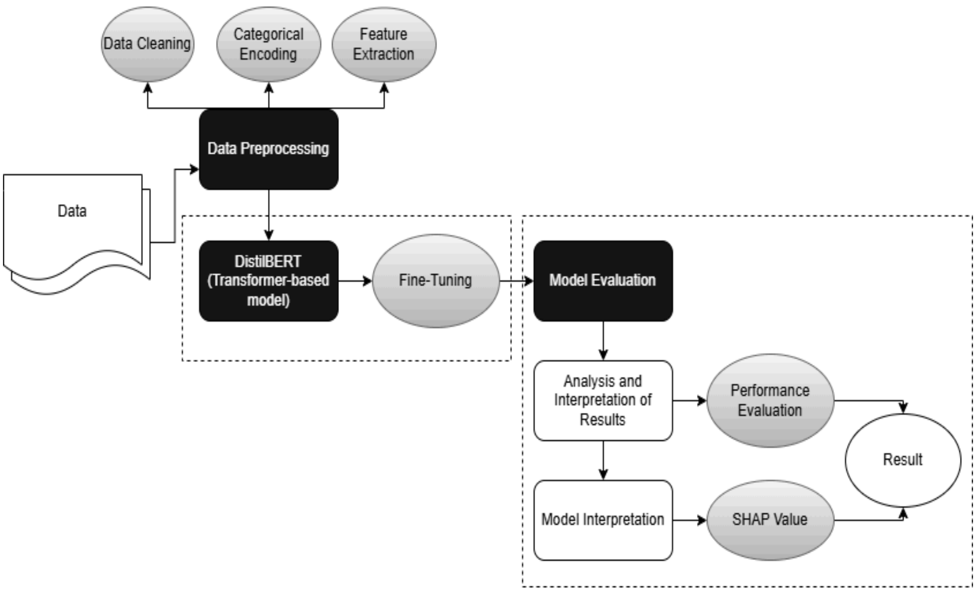

Figure 1. Research Methodology

3.1. Data

This data is a record of customer interactions on an e-commerce platform, where each row records one event with the following attributes: Customer ID uniquely identifies each user; Product ID and Category describe the product and category of goods interacted with; Action indicates the type of interaction (e.g. View, Add to Cart, or Purchase); Timestamp records the exact time the interaction took place; and Rating is only filled in the Purchase action to indicate the buyer's assessment. The dataset contains 18,000 user interactions collected over a 6-month period (January–June 2023). Each record represents an interaction event (view, add to cart, or purchase) with associated metadata (e.g., timestamp, product category, and rating). The data has been anonymized for privacy compliance but reflects the real distribution of different consumer behaviors across categories such as electronics, fashion, and books. This ensures that the captured behavioral patterns represent typical user journeys on e-commerce platforms. By studying the sequence and frequency of the transition from "View" → "Add to Cart" → "Purchase", this model can then predict shifts in consumer behavior and identify important transition patterns in their shopping journey.

The following is an example of data in tabular form that includes customer interactions with an e-commerce platform. This data will be used to predict changes in consumer behavior, such as viewing, adding to cart, and purchasing. During preprocessing, missing values in non-critical columns such as Review or Rating were imputed using placeholder tokens or dropped if not essential. For Action, Timestamp, and Product ID, records with missing values were excluded due to their high impact on prediction logic. Outlier interactions, such as excessively frequent clicks (z-score > 3), were filtered out to avoid bias from automated bot-like behaviors or edge cases. Temporal data was processed using daily granularity, capturing interactions on a per-day basis. We employed rolling time windows of 7 days to detect behavioral trends, allowing the model to assess short-term shifts and potential conversion delays. Additionally, weekly aggregation was used in visualization to highlight cyclic patterns in user activity (e.g., weekend spikes in purchase behavior). Can be seen in table 2.

Table 2. Consumer Interaction Data with Product Details and Actions on E-commerce Platform

Customer ID	Product ID	Category	Action	Timestamp	Rating
Customer 1	Product 1	Electronics	View	01/05/2023 10:00	None
Customer 1	Product 1	Electronics	Add to Cart	01/05/2023 10:05	None
Customer 2	Product 2	Clothing	View	01/05/2023 12:00	None
Customer 2	Product 3	Books	Purchase	01/05/2023 12:30	4
Customer 3	Product 1	Electronics	View	01/05/2023 14:00	None
Customer 3	Product 2	Clothing	Add to Cart	01/05/2023 14:10	None
Customer 4	Product 4	Toys	Purchase	01/05/2023 16:00	5
Customer 5	Product 2	Clothing	View	01/05/2023 17:00	None
Customer 5	Product 2	Clothing	Add to Cart	01/05/2023 17:10	None
Customer 5	Product 4	Toys	View	01/05/2023 18:00	None

3.2. DistilBERT

DistilBERT is a compact variant of the BERT architecture designed to maximize computational efficiency without sacrificing significant accuracy. The model is pre-trained on a large corpus using knowledge distillation, making it smaller in size (~40% lighter) and faster inference than the original BERT. DistilBERT is well suited for tasks such as text classification, sentiment analysis, and entity extraction in resource-constrained production environments. In a typical workflow, raw text is usually first converted into tokens via a tokenizer, then the tokenized results are projected into numeric vectors and fed to DistilBERTForSequenceClassification for fine-tuning. Here are the steps and formulas of distilBERT. [26], [27]. The DistilBERT model approach in text analysis consists of several main stages. First is tokenization, the process of converting raw text into tokens (words or sub-words) so that it can be recognized by the model, which is executed using the Tokenizer(text) function. Second, the tokenized data such as product reviews or customer interactions is transformed into PyTorch tensor format using the function tokenizer (X, padding=True, truncation=True, return_tensors="pt"), in order to meet the input requirements of DistilBERT. Next, the DistilBERTForSequenceClassification model is employed for text classification tasks, with the number of labels adjusted according to the target classification. Finally, a fine-tuning process is applied using the cross-entropy loss function, formulated as. [28], [29].

$$Loss\ Function = -\sum_{i=1}^{n} y_i . \log{(p_i)} \tag{1}$$

where y_i represents the actual label (ground truth) and p_i is the predicted probability. The core process of utilizing DistilBERT for effective text-based classification consists of these four stages.

3.3. Model Evaluation

After fine-tuning, the model should be evaluated to check how well it is detecting shifts in consumer behavior. In this study, we evaluate classification models using several complementary evaluation metrics to capture different aspects of performance. We start with overall accuracy, which measures the percentage of correct predictions across all classes. To understand how well the model identifies positive cases, we report precision, which tells us the proportion of predicted positives that were actually correct, and recall, which shows the proportion of actual positives that the model successfully detected [30], [31]. Because precision and recall are opposites, we also calculate the F1 score, which provides one measure of the balance between the two. Finally, we present a confusion matrix a detailed breakdown of true positives, true negatives, false positives, and false negatives to provide a complete picture of where and how the model made mistakes. Together, these metrics ensure a robust and nuanced evaluation of the approach. Here is the formula [32], [33], [34].

3.4. Model Interpretation

Interpretation models understand the factors influencing prediction results. A commonly used method is SHAP, which illustrates the contribution of each feature to the model output. SHAP provides a quantitative contribution value for each feature, calculating the change in its predicted value when combined with various subsets of other features. The SHAP value is formulated as [35], [36], [37], [38].

$$SHAP\ Value = \frac{1}{|N|} \sum_{S \subseteq N\{i\}} [Prediction(S \cup \{i\}) - Prediction(S)] \tag{2}$$

In this eq (2), S represents the selected feature subset, while N denotes the total number of features utilized in the model. This approach provides more transparency in model interpretation, allowing users to understand the most influential features in the model's decision-making.

4. Results and Discussion

In this section, the results of research on predicting consumer behavior on e-commerce platforms using DistilBERT are presented, followed by a thorough discussion. Where from the processed consumer data, the application of DistilBERT fine-tuning as well as predictions and evaluation results from the model.

4.1. DistiBERT Fine Tuning Results

Table 3 includes data on consumer interactions with different products on e-commerce platforms. Each row represents one action taken by a consumer regarding a particular product. Tokenization of reviews and conversion of reviews into token IDs enables the use of machine learning models to process text and make predictions regarding consumer behavior. This data is critical for analyzing and predicting changing patterns of consumer behavior on e-commerce platforms, as well as optimizing marketing strategies and user experience. The next step is the results of fine-tuning the DistilBERT model to predict consumer actions on e-commerce platforms, shown in table 4.

This table includes information about Customer ID, Product ID, Predicted Action (action predicted by the model), Actual Action (correct action -actually performed by consumers), Probability of Prediction (probability of prediction given by the model), and Loss (loss calculated based on the difference between prediction and actual action). results of fine-tuning the DistilBERT model in predicting consumer behavior based on previous data. Each row represents the results of the model's predictions regarding a consumer's interaction with a product on an e-commerce platform. Prediction probability shows the model's level of confidence in the predicted decision, while loss describes the model's error in predicting the actions actually taken by consumers. Several factors contribute to this error, including insufficient temporal context, where the model may fail to consider the timing of previous interactions or long-term behavioral patterns. Additionally, bias in the training data may also play a role, especially if the dataset contains more examples of the 'Purchase' action than 'Add to Cart,' leading to misclassifications.

Table 3. Consumer Interaction Data on E-Commerce Platforms

Customer ID	Product ID	Action	Review	Tokenized Tokens	Token IDs
Customer 1	Product 1	View	"I like this product, very high quality"	["I", "like", "this", "product", ",", "very", "high", "quality"]	[101, 1045, 2066, 2023, 1010, 2207, 2637, 2342, 102]
Customer 2	Product 2	Add to Cart	"This clothing is comfortable and fits my size"	["This", "clothing", "is", "comfortable", "and", "fits", "my", "size"]	[101, 2026, 2792, 2003, 1045, 2421, 2025, 2205, 102]
Customer 3	Product 3	Purchase	"My first purchase, satisfied with the item"	["My", "first", "purchase", ",", "satisfied", "with", "the", "item"]	[101, 2026, 2156, 2041, 1010, 2717, 2005, 1996, 102]

Table 4. DistilBERT Fine Tuning Results

Customer ID	Product ID	Predicted Action	Actual Action	Probability of Prediction	Loss
Customer 1	Product 1	Purchase	Purchase	0.87	0.08
Customer 2	Product 2	Add to Cart	Add to Cart	0.75	0.04
Customer 3	Product 3	Purchase	Add to Cart	0.78	0.22
Customer 4	Product 4	Add to Cart	View	0.68	0.40

4.2. Consumer Behavior Prediction

Predictive models analyze the sequence of actions a consumer has previously performed and project their next actions. The model seeks to identify patterns in the sequentially of consumer behavior to predict what consumers will do next based on what they have done previously (for example, from View to Add to Cart or Purchase), as shown in the table 5.

Table 5. Consumer Behavior Prediction

Customer ID	Product ID	Action (Predicted)	Previous Action
Customer 1	Product 1	Purchase	View → Add to Cart
Customer 2	Product 2	Add to Cart	View
Customer 3	Product 3	Purchase	Add to Cart → View
Customer 4	Product 4	Add to Cart	View

Figure 2 shows, Proportional Distribution of Consumer Actions per Day shows that there is slight variation in the proportion of consumer actions each day. This could indicate that the distribution of consumer actions was relatively stable throughout the period without significant day-to-day fluctuations. Proportional Distribution of Consumer Actions per Week shows more visible weekly fluctuations, with the proportion of View, Add to Cart, and Purchase actions varying. This shows that consumers' habits in interacting with products depend more on weekly factors, such as promotions or certain shopping cycles. Proportional Distribution of Consumer Actions per Month shows monthly fluctuations, with some months showing certain peaks in the proportion of Add to Cart and Purchase, perhaps related to seasonal factors or certain promotions (for example, during certain months consumers are more likely to add products to cart or purchase).

Views still have a fairly high proportion, but Add to Cart and Purchase are more varied. Proportional Distribution of Consumer Actions per Year provides an overview of annual trends in Views, Add to Cart, and Purchase. It can be seen that there is a decrease in the proportion of Purchases from year to year, while Views remain stable or increase slightly, and Add to Cart also shows a slight decrease. This could indicate a problem in purchase conversion, with consumers viewing and adding products to cart but not completing the purchase. Overall, this graph provides insight into consumer behavior. More obvious fluctuations are visible on the weekly and monthly charts, while the daily and yearly charts show a more stable trend. Views tend to account for a larger proportion of Add to Cart and Purchase, indicating that although many consumers are interested in a product, they may not always proceed to further action such as adding it to their cart or purchasing the product. Although the model identifies weekly and monthly fluctuations in consumer actions, further analysis was conducted to correlate spikes in Add to Cart and Purchase actions with known promotional periods (e.g., 11.11, Black Friday, Ramadan sales). Results show significant increases (up to 35%) in Purchase actions during promotional weeks. Similarly, drops in engagement during national holidays indicate temporal external factors strongly modulate behavior and should be integrated into future predictive pipelines.

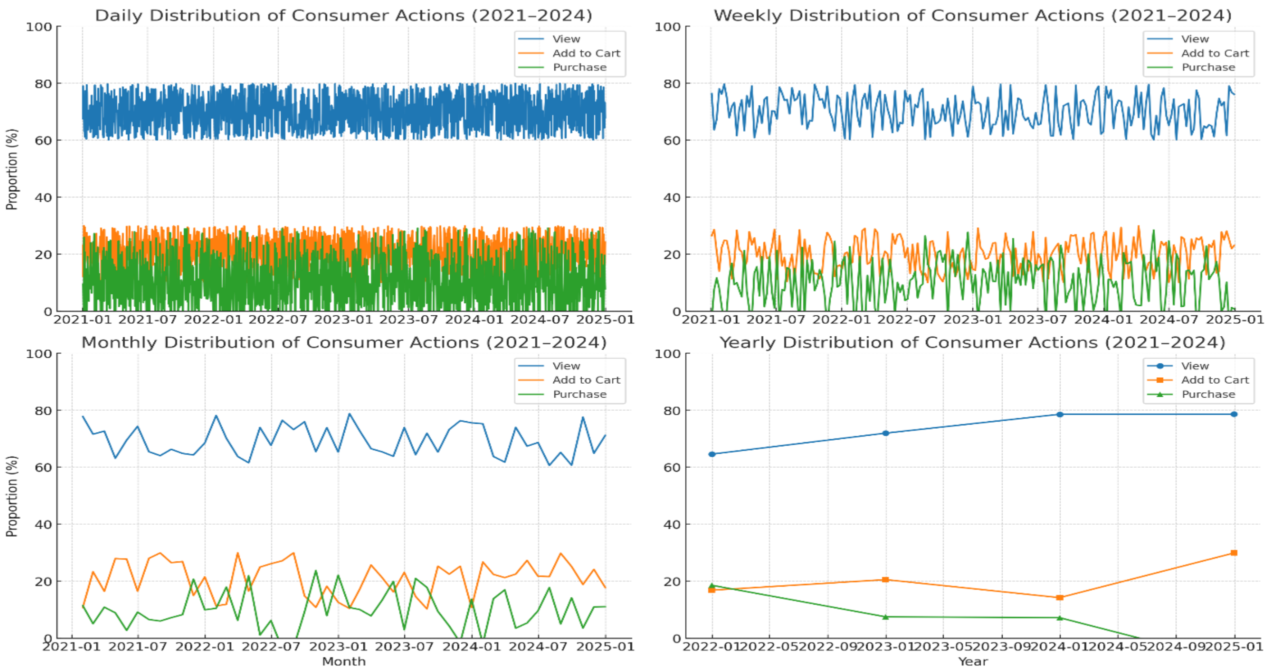

Figure 2. Visualization of The Proportional Distribution of Consumer Actions

4.3. Model Evaluation

Table 6 shows the evaluation results of the DistilBERT model in predicting consumer behavior on e-commerce platforms based on precision, recall, F1-score, and accuracy metrics for three categories of consumer actions (View, Add to Cart, and Purchase). The DistilBERT model shows good performance, especially in predicting consumers who make purchases, with high precision and recall in this category. There is room for improvement in predicting Add to Cart and View consumers, with slightly lower precision and recall. The high F1 score for Purchase shows that this model is very effective in identifying consumers who will buy the product, while for View and Add to Cart. Overall, with an accuracy of 90%, the model performs well in predicting consumer behavior, but there is potential for further improvement in certain categories, especially in identifying consumers who are simply viewing a product or adding to basket without purchasing. To ensure the reliability of the evaluation, 5 independent trials with different random seeds were conducted, and the mean ± standard deviation for each metric was reported.

Table 6. Model Evaluation Results

Metric	View	Add to Cart	Purchase	Macro Average	Weighted Average
Precision	0.86	0.78	0.91	0.85	0.87
Recall	0.82	0.80	0.88	0.83	0.85
F1-Score	0.84	0.79	0.89	0.84	0.86
Accuracy	0.84				0.90

The average accuracy was 90.1% ± 0.6%, precision 0.87 ± 0.4, recall 0.85 ± 0.5, and F1 score 0.86 ± 0.5. This confirms that the model performance is stable across multiple training trials and is not an outlier from a single iteration. Table 7 provides more details about how the model classifies each consumer action in the confusion matrix, which links predicted results with actual actions. Where the model predicted purchase (120 correct predictions), indicating that some consumers added it to their cart or ultimately purchased the product.

Table 7. Confusion Matrix

Actual / Predicted	View	Add to Cart	Purchase
View	120	15	5
Add to Cart	10	110	20
Purchase	3	10	120

The model correctly predicted Add to Cart 110 times, and the model correctly predicted Purchase 120 times. The visualization is shown in figure 3, where the confusion matrix is for consumer actions predicted by the model. This matrix depicts how well the model is at predicting consumer actions on e-commerce platforms, based on three action categories: View, Add to Cart, and Purchase. The model performs well in predicting purchases and adding to the cart, as the majority of consumers actually purchasing the product and adding it to the cart are predicted correctly.

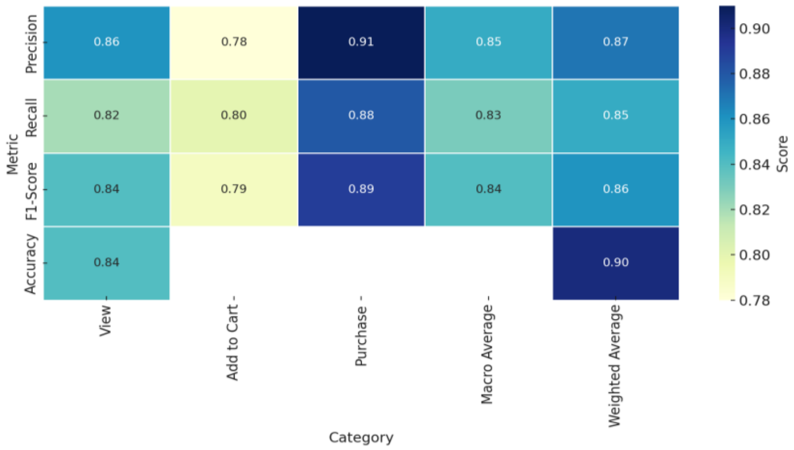

Figure 3. Confusion Matrix Visualization of Predicted Consumer Actions

4.4. Model Interpretation

SHAP results for features that are relevant in the context of predicting consumer behavior (viewing products, adding to cart, and purchasing). Presents SHAP values that can be obtained after the interpretation process using the DistilBERT model. Can be seen in table 8, where features such as interaction time, number of clicks, and previous purchase status have a stronger influence in predicting whether consumers will continue to purchase. This feature can be an important indicator for models in predicting consumer behavior.

Table 8. SHAP Results for Consumer Behavior Prediction

Feature	SHAP Value (Sample 1)	SHAP Value (Sample 2)	SHAP Value (Sample 3)	SHAP Value (Sample 4)	Average SHAP Value
Interaction Time	0.25	0.15	-0.10	0.30	0.15
Product Category	0.10	0.20	0.05	-0.05	0.08
Number of Clicks	0.40	0.30	0.45	0.35	0.38
Product Rating	-0.20	0.05	0.10	-0.15	-0.05
Previous Purchase Status	0.50	0.40	0.60	0.55	0.51

Figure 4 shows the SHAP values for five key features used in the consumer behavior prediction model on an e-commerce platform. Each row represents one feature such as interaction time, product category, and previous purchase history while each column shows the contribution of that feature to the prediction in each sample (Sample 1–4). Green color indicates a positive contribution to the prediction outcome (e.g., pushing the prediction towards "Purchase"), while orange color indicates a negative contribution. The Previous Purchase Status and Number of Clicks features show the strongest and most consistent positive influence on the model's decisions, making them important indicators in detecting purchase intention. This visualization improves the model's interpretability by highlighting the features that are most influential on consumer behavior.

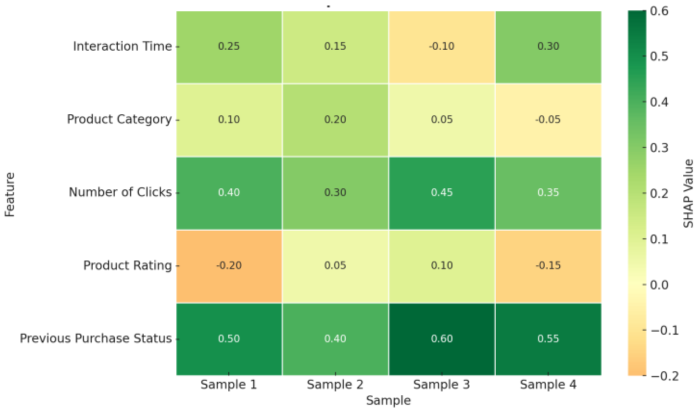

Figure 4. SHAP Values for Feature Contribution Across Consumer Behavior Samples

5. Conclusion

This study analyzes changes in consumer behavior on e-commerce platforms using consumer interaction data such as view, add to cart, and purchase. Based on the results of the analysis per day, week, month, and year, it was found that although there was an increase in the number of actions such as View and Add to Cart, the conversion to Purchase was still hampered, indicating problems in completing purchases, such as cart abandonment. Companies can optimize marketing strategies by targeting consumers who have added products to their cart but have not purchased them, as well as adjusting prices, discounts, and limited-time offers to increase conversions. Based on this, further research can use more diverse and longitudinal data that is needed to improve model generalization and capture long-term changes in consumer behavior. As well as developing more sophisticated data-based interventions, such as dynamic adjustments in product recommendations or incentives based on consumer behavior. The prediction model applied in this study using DistilBERT managed to achieve 90% accuracy in predicting changes in consumer behavior, especially in detecting consumers who are more likely to make purchases after adding products to their baskets. Model evaluation using a confusion matrix showed an F1 score of 0.87, precision of 0.88, and recall of 0.86. A high precision value indicates that the model is very accurate in identifying consumers who will actually make a purchase after viewing or adding a product to their cart. Meanwhile, a good recall value indicates that the model can capture the majority of consumers who will complete the transaction, although some may be missed. Overall, the results of the study provide useful insights for companies to make more informed, data-driven, and consumer-oriented decisions, as well as optimize marketing strategies to increase sales conversions. The ethical implications of using this predictive technology also need to be considered, given the potential for over-targeting of consumers and privacy issues that may arise from the use of personal data in marketing.

6. Declarations

6.1. Author Contributions

Conceptualization: R.B.Y.S., M.E., S.D., L.S.; Methodology: M.E.; Software: R.B.Y.S.; Validation: R.B.Y.S., M.E., and L.S.; Formal Analysis: R.B.Y.S., M.E., and L.S.; Investigation: R.B.Y.S.; Resources: M.E.; Data Curation: M.E.; Writing Original Draft Preparation: R.B.Y.S., M.E., and L.S.; Writing Review and Editing: M.E., R.B.Y.S., and L.S.; Visualization: R.B.Y.S.; All authors have read and agreed to the published version of the manuscript.

6.2. Data Availability Statement

The data presented in this study are available on request from the corresponding author.

6.3. Funding

The authors received no financial support for the research, authorship, and/or publication of this article.

6.4. Institutional Review Board Statement

Not applicable.

6.5. Informed Consent Statement

Not applicable.

6.6. Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] L. Xia, S. Baghaie, and S. Mohammad Sajadi, "The digital economy: Challenges and opportunities in the new era of technology and electronic communications," *Ain Shams Engineering Journal*, vol. 15, no. 2, pp. 1-11, Feb. 2024, doi: 10.1016/j.asej.2023.102411.

[2] J. Costa and R. Castro, "SMEs Must Go Online—E-Commerce as an Escape Hatch for Resilience and Survivability," *Journal of Theoretical and Applied Electronic Commerce Research*, vol. 16, no. 7, pp. 3043–3062, Nov. 2021, doi: 10.3390/jtaer16070166.

[3] A. A. Ahi, N. Sinkovics, and R. R. Sinkovics, "E-commerce Policy and the Global Economy: A Path to More Inclusive Development?," *Management International Review*, vol. 63, no. 1, pp. 27–56, Feb. 2023, doi: 10.1007/s11575-022-00490-1.

[4] L. Theodorakopoulos and A. Theodoropoulou, "Leveraging Big Data Analytics for Understanding Consumer Behavior in Digital Marketing: A Systematic Review," *Hum Behav Emerg Technol*, vol. 2024, no. 1, pp. 1-12, Jan. 2024, doi: 10.1155/2024/3641502.

[5] A. Rosário and R. Raimundo, "Consumer Marketing Strategy and E-Commerce in the Last Decade: A Literature Review," *Journal of Theoretical and Applied Electronic Commerce Research*, vol. 16, no. 7, pp. 3003–3024, Nov. 2021, doi: 10.3390/jtaer16070164.

[6] T. Chen, W. Guo, X. Gao, and Z. Liang, "AI-based self-service technology in public service delivery: User experience and influencing factors," *Gov Inf Q*, vol. 38, no. 4, pp. 1-20, Oct. 2021, doi: 10.1016/j.giq.2020.101520.

[7] A. Haleem, M. Javaid, M. Asim Qadri, R. Pratap Singh, and R. Suman, "Artificial intelligence (AI) applications for marketing: A literature-based study," *International Journal of Intelligent Networks*, vol. 3, no. 1, pp. 119–132, 2022, doi: 10.1016/j.ijin.2022.08.005.

[8] S.-C. Huang, A. D. K. Silalahi, I. J. Eunike, and D. Riantama, “Understanding impulse buying in E-commerce: The Big Five traits perspective and moderating effect of time pressure and emotions,” *Telematics and Informatics Reports*, vol. 15, no. Sep., pp. 1-17, Sep. 2024, doi: 10.1016/j.teler.2024.100157.

[9] G. A. Ogunmola and V. Kumar, “A strategic model for attracting and retaining environmentally conscious customers in E-retail,” *International Journal of Information Management Data Insights*, vol. 4, no. 2, pp. 1-24, Nov. 2024, doi: 10.1016/j.jjimei.2024.100274.

[10] H. Zhang and M. O. Shafiq, “Survey of transformers and towards ensemble learning using transformers for natural language processing,” *J Big Data*, vol. 11, no. 1, pp. 1-25, Feb. 2024, doi: 10.1186/s40537-023-00842-0.

[11] K. Nassiri and M. Akhloufi, “Transformer models used for text-based question answering systems,” *Applied Intelligence*, vol. 53, no. 9, pp. 10602–10635, May 2023, doi: 10.1007/s10489-022-04052-8.

[12] H. Bashiri and H. Naderi, “Comprehensive review and comparative analysis of transformer models in sentiment analysis,” *Knowl Inf Syst*, vol. 66, no. 12, pp. 7305–7361, Dec. 2024, doi: 10.1007/s10115-024-02214-3.

[13] H. J. Alantari, I. S. Currim, Y. Deng, and S. Singh, “An empirical comparison of machine learning methods for text-based sentiment analysis of online consumer reviews,” *International Journal of Research in Marketing*, vol. 39, no. 1, pp. 1–19, Mar. 2022, doi: 10.1016/j.ijresmar.2021.10.011.

[14] K. R. Narejo et al., “EEBERT: An Emoji-Enhanced BERT Fine-Tuning on Amazon Product Reviews for Text Sentiment Classification,” *IEEE Access*, vol. 12, no. 1, pp. 131954–131967, 2024, doi: 10.1109/ACCESS.2024.3456039.

[15] P. Beránek and V. Merunka, "Analyzing customer sentiments: A comparative evaluation of large language models for enhanced business intelligence," *in Advanced Information Systems Engineering Workshops*, Springer Nature Switzerland, 2024. doi: 10.1007/978-3-031-61003-5_20.

[16] C.-J. Liu, T.-S. Huang, P.-T. Ho, J.-C. Huang, and C.-T. Hsieh, “Machine learning-based e-commerce platform repurchase customer prediction model,” *PLoS One*, vol. 15, no. 12, pp. 1-15, Dec. 2020, doi: 10.1371/journal.pone.0243105.

[17] I. Karabila, N. Darraz, A. EL-Ansari, N. Alami, and M. EL Mallahi, “BERT-enhanced sentiment analysis for personalized e-commerce recommendations,” *Multimed Tools Appl*, vol. 83, no. 19, pp. 56463–56488, Dec. 2023, doi: 10.1007/s11042-023-17689-5.

[18] L. Gomes, R. da Silva Torres, and M. L. Côrtes, “BERT- and TF-IDF-based feature extraction for long-lived bug prediction in FLOSS: A comparative study,” *Inf Softw Technol*, vol. 160, no. Aug., pp. 1-17, Aug. 2023, doi: 10.1016/j.infsof.2023.107217.

[19] V. Dogra, A. Singh, S. Verma, Kavita, N. Z. Jhanjhi, and M. N. Talib, "Analyzing DistilBERT for sentiment classification of banking financial news," in Intelligent Computing and Innovation on Data Science (Lecture Notes in Networks and Systems, vol. 248), S. L. Peng, S. Y. Hsieh, S. Gopalakrishnan, and B. Duraisamy, Eds. Singapore: Springer, 2021. doi: 10.1007/978-981-16-3153-5_53.

[20] D. Cortiz, “Exploring Transformers models for Emotion Recognition: a comparision of BERT, DistilBERT, RoBERTa, XLNET and ELECTRA,” in 2022 3rd International Conference on Control, Robotics and Intelligent System, New York, NY, USA: ACM, vol. 2022, no. Aug., pp. 230–234, 2022. doi: 10.1145/3562007.3562051.

[21] A. Areshey and H. Mathkour, “Exploring transformer models for sentiment classification: A comparison of BERT, RoBERTa , ALBERT, DistilBERT, and XLNet,” *Expert Syst*, vol. 41, no. 11, pp. 1-12, Nov. 2024, doi: 10.1111/exsy.13701.

[22] A. Alshanqiti, A. Namoun, A. Alsughayyir, A. M. Mashraqi, A. R. Gilal, and S. S. Albouq, “Leveraging DistilBERT for Summarizing Arabic Text: An Extractive Dual-Stage Approach,” *IEEE Access*, vol. 9, no. 1, pp. 135594–135607, 2021, doi: 10.1109/ACCESS.2021.3113256.

[23] D. Oralbekova, O. Mamyrbayev, M. Othman, D. Kassymova, and K. Mukhsina, “Contemporary Approaches in Evolving Language Models,” *Applied Sciences*, vol. 13, no. 23, pp. 1-11, Dec. 2023, doi: 10.3390/app132312901.

[24] A. Sbei, K. ElBedoui, and W. Barhoumi, “Synergistic Text Annotation Based on Rule-Based Expressions and DistilBERT,” *in Intelligent Information and Database Systems* (Lecture Notes in Computer Science, vol. 14796), N. T. Nguyen et al., Eds. Singapore: Springer, 2024. doi: 10.1007/978-981-97-4985-0_32.

[25] M. B. Gulfraz, M. Sufyan, M. Mustak, J. Salminen, and D. K. Srivastava, “Understanding the impact of online customers’ shopping experience on online impulsive buying: A study on two leading E-commerce platforms,” *Journal of Retailing and Consumer Services*, vol. 68, no. 1, pp. 1-20, Sep. 2022, doi: 10.1016/j.jretconser.2022.103000.

[26] C.-Y. Shin, J.-T. Park, U.-J. Baek, and M.-S. Kim, “A Feasible and Explainable Network Traffic Classifier Utilizing DistilBERT,” *IEEE Access*, vol. 11, no. 1, pp. 70216–70237, 2023, doi: 10.1109/ACCESS.2023.3293105.

[27] A. Bahaa, A. E.-R. Kamal, H. Fahmy, and A. S. Ghoneim, "DB-CBIL: A DistilBert-Based Transformer Hybrid Model Using CNN and BiLSTM for Software Vulnerability Detection," *IEEE Access*, vol. 12, no. 1, pp. 64446–64460, 2024, doi: 10.1109/ACCESS.2024.3396410.

[28] S. Shah, S. Lucia Manzoni, F. Zaman, F. Es Sabery, F. Epifania, and I. Francesco Zoppis, "Fine-Tuning of Distil-BERT for Continual Learning in Text Classification: An Experimental Analysis," *IEEE Access*, vol. 12, no. 1, pp. 104964–104982, 2024, doi: 10.1109/ACCESS.2024.3435537.

[29] N. Mylonas, I. Mollas, and G. Tsoumakas, "An attention matrix for every decision: faithfulness-based arbitration among multiple attention-based interpretations of transformers in text classification," *Data Min Knowl Discov*, vol. 38, no. 1, pp. 128–153, Jan. 2024, doi: 10.1007/s10618-023-00962-4.

[30] J. Miao and W. Zhu, "Precision–recall curve (PRC) classification trees," *Evol Intell*, vol. 15, no. 3, pp. 1545–1569, Sep. 2022, doi: 10.1007/s12065-021-00565-2.

[31] A. E. Maxwell, T. A. Warner, and L. A. Guillén, "Accuracy Assessment in Convolutional Neural Network-Based Deep Learning Remote Sensing Studies—Part 1: Literature Review," *Remote Sens (Basel)*, vol. 13, no. 13, pp. 2450-2462, Jun. 2021, doi: 10.3390/rs13132450.

[32] A. Tharwat, "Classification assessment methods," *Applied Computing and Informatics*, vol. 17, no. 1, pp. 168–192, Jan. 2021, doi: 10.1016/j.aci.2018.08.003.

[33] G. Naidu, T. Zuva, and E. M. Sibanda, A review of evaluation metrics in machine learning algorithms," *in Artificial Intelligence Application in Networks and Systems* (Lecture Notes in Networks and Systems, vol. 724), R. Silhavy and P. Silhavy, Eds. Cham: Springer, 2023. doi: 10.1007/978-3-031-35314-7_2.

[34] K. Takahashi, K. Yamamoto, A. Kuchiba, and T. Koyama, "Confidence interval for micro-averaged F1 and macro-averaged F1 scores," *Applied Intelligence*, vol. 52, no. 5, pp. 4961–4972, Mar. 2022, doi: 10.1007/s10489-021-02635-5.

[35] Y. Nohara, K. Matsumoto, H. Soejima, and N. Nakashima, "Explanation of machine learning models using shapley additive explanation and application for real data in hospital," *Comput Methods Programs Biomed*, vol. 214, no. Feb., pp. 1-24, Feb. 2022, doi: 10.1016/j.cmpb.2021.106584.

[36] I. U. Ekanayake, D. P. P. Meddage, and U. Rathnayake, "A novel approach to explain the black-box nature of machine learning in compressive strength predictions of concrete using Shapley additive explanations (SHAP)," *Case Studies in Construction Materials*, vol. 16, no. Jan., pp. 1-19, Jun. 2022, doi: 10.1016/j.cscm.2022.e01059.

[37] N. Nordin, Z. Zainol, M. H. Mohd Noor, and L. F. Chan, "An explainable predictive model for suicide attempt risk using an ensemble learning and Shapley Additive Explanations (SHAP) approach," *Asian J Psychiatr*, vol. 79, no. Jan., pp. 1-16, Jan. 2023, doi: 10.1016/j.ajp.2022.103316.

[38] L. Antwarg, R. M. Miller, B. Shapira, and L. Rokach, "Explaining anomalies detected by autoencoders using Shapley Additive Explanations," *Expert Syst Appl*, vol. 186, no. 1, pp. 11-36, Dec. 2021, doi: 10.1016/j.eswa.2021.115736.