

# A Machine Learning Framework for Predicting Online Purchasing Intensity Using Composite Behavioral Indices and Balanced Classification Techniques

L. Anju<sup>1\*</sup>, and Dr.S. Veni<sup>2</sup>

<sup>1\*</sup>Research Scholar, Department of Computer Science, Karpagam Academy of Higher Education (Deemed to be University), Coimbatore, Tamil Nadu, India. anjulekha2026@gmail.com, <https://orcid.org/0009-0007-2412-8913>

<sup>2</sup>Professor (Research Guide), Department of Computer Science, Karpagam Academy of Higher Education (Deemed to be University), Coimbatore, Tamil Nadu, India. venics@kahedu.edu.in, <https://orcid.org/0000-0002-2999-8463>

Received: September 25, 2025; Revised: October 31, 2025; Accepted: December 24, 2025; Published: February 27, 2026

## Abstract

The problem of correctly predicting the intensity of online purchasing proves still to be very challenging due to scattered behavioural cues, extreme class imbalance and the increasing importance of trust and security in the decision-making process of consumers. Purchase prediction models currently in use have mostly been based on single behavioural measures, i.e. click rate, purchase history, etc. Single behavioural measures are not sufficient to capture the multidimensionality of purchase intention in actual e-commerce settings. To overcome these drawbacks in current paper proposes the development of Security-Conscious Composite Behavioural Intelligence (SC-CBI) model for the forecasting of online purchasing intensity at multiple levels. The proposed model introduces Composite Behavioural Indices (CBI) that introduce behavioural engagement, temporal consistency, monetary commitment and security-induced trust interactions in a low dimensional representation of user intent. In addition, a Security-Aware Balanced Classification Strategy (SABCS) will be developed in order to overcome the severe class imbalance and to preserve trust-related behavioural distributions by adaptive resampling and risk-sensitive cost optimisation. Tests on large-scale, real-world e-commerce data show that the given framework is better than traditional feature-based and imbalance-blind models. The SC-CBI model has an average accuracy of 92.7, a macro F1-score of 90.8 and a balanced recall of 90.2 which can improve up to 8-12 points compared to baselines from strong ensembles. The given framework is most appropriate for implementation in the contemporary e-commerce systems that require precision decision support in the security impaired and data imbalanced environment.

**Keywords:** Online Purchasing Intensity, Composite Behavioural Index, Security-Aware Learning, Class Imbalance, Trust Modelling, E-Commerce Analytics, Machine Learning, Ensemble Classification.

## 1 Introduction

Over the past few years, the concept of digital commerce has radically altered the nature of consumer-to-consumer online interaction by generating vast amounts of high-frequency behavioural data, such as browsing, cart, purchase, and security behaviour prompted by transactions leading to payments.

The importance of security and trust as major determinants of online purchasing behaviour has been acknowledged to be growing (Wei et al., 2022). Authentication, payment verification, failed payments, and dispute resolution are important user experiences that determine purchasing confidence and the intensity of future engagement (Handoyo, 2024). Nonetheless, the current models of purchasing intensity prediction hardly take into account the interaction signals related to security systematically (Zhou & Hudin, 2024). The lack of this decreases the interpretability and robustness of the models, especially in those environments where the issue of fraud prevention, cyber threats, and regulatory compliance is essential.

This paper addresses these problems by proposing a new machine learning model for predicting online purchasing intensity using Composite Behavioural Indices (CBI) and a Security-Aware Balanced Classification Strategy. The framework grouping the engagements, temporal stability, monetary investment, and trust-based interaction together into a single behavioural representation facilitates the learning process in a stable and interpretable manner. Also, a balancing mechanism that ensures security considers trust-aware resampling, and a risk-sensitive cost function is used to address class imbalance without altering behavioural patterns. The hypothesised model is confirmed using massive real-world e-commerce data, and it can be easily integrated into contemporary digital commerce systems.

The rest of this paper is structured in the following way. Section 2 conducts a literature review of online purchase behaviour modelling, machine learning, managing class imbalance and security-aware analytics. Section 3 describes the proposed system architecture. Section 4 details the dataset and experimental setup. Section 5 introduces the Composite Behavioural Index design, while Section 6 presents the Security-Aware Balanced Classification Strategy along with the proposed algorithm. Section 7 discusses the machine learning models and ensemble learning strategy. Section 8 reports experimental results, robustness analysis, and ablation studies. Section 9 discusses deployment and security implications, and Section 10 concludes the paper.

## 2 Related Work and Research Positioning

The study of online purchase behaviour modelling has taken shape from simple transactional analysis to sophisticated behavioural analytics models. Initial research was mainly based on non-dynamic measures, such as purchase frequency, recency, and monetary metrics, to segment people and estimate purchasing behaviour (Wei et al., 2022; Ullah et al., 2023). Future studies added clickstream information, browsing time, and cart abandonment to gain more insight into behavioural data (Kukar-Kinney et al., 2022; Kontola, 2024). Recently, sequential and temporal modelling, such as Markov chains and recurrent neural networks, have been proposed to model behavioural evolution over time.

Additionally, the majority of models of behaviour currently available give much attention to engagement or transaction frequency but ignore other important dimensions such as temporal consistency, monetary commitment, and trust-related interactions (Handoyo, 2024; Naeem, 2025). E-commerce analytics has also actively embraced machine learning methods for recommender systems, demand prediction, churn, and purchase probability estimation (Dritsas & Trigka, 2025). Classical supervised models such as logistic regression, decision trees and support vector machines are used

because they are easy to use and interpret. Ensemble and deep learning methods are more recent, including random forests, gradient boosting, and attention-based neural networks, which have been reported to exhibit better predictive accuracy due to their ability to capture nonlinear relationships.

Besides, high-dimensional raw feature spaces are the basis of most models and are susceptible to sparsity and behavioural noise. Purchasing prediction should also be critical on the imbalance of the classes distribution as the extremely large proportion of users is not buying or buying at high rates, and high-intensity purchasers are a minority but economically significant (Wang et al., 2021). Conventional methods of balancing (random oversampling, undersampling, synthetic sample generation, e.g. SMOTE) may frequently corrupt behavioural distributions and produce artificial samples (Chen et al., 2025). Such problems can be partly mitigated by use of cost-sensitive learning, however contextual risk and trust is usually ignored. Most imbalance-handling approaches are not based on domain specific aspects like perceived risk, system reliability, or security behaviour, which makes them less applicable in a real-world setting (Martins, 2023). Security and trust have turned out to be the major elements of online purchasing behaviour, particularly when it comes to dealing with money and other personal information that is sensitive (Saeed, 2023).

The studies have shown that authentication systems, the outcome of payment verification and transaction failures affect the confidence of the users and future purchasing behaviour (Alabi, 2022). However, the security-related interactions are mostly researched separately to identify the fraud and not within other behavioural prediction models (Rofi'i, 2023). This split compromises the ability to estimate long-term buying power and fuels the urge to learn with regard to security. Purchasing intensity prediction is a hybrid learning and behavioural and security-seeking problem in comparison to the current methods as proposed by the study. The given structure capitalises on involvement, time regularity, financial investment, and trust relations in Composite Behavioural Indices. It is a noise-resistant interpretable context-sensitive solution, as it combines them with a Security-Aware Balanced Classification Strategy.

### 3 System Overview and Framework Architecture

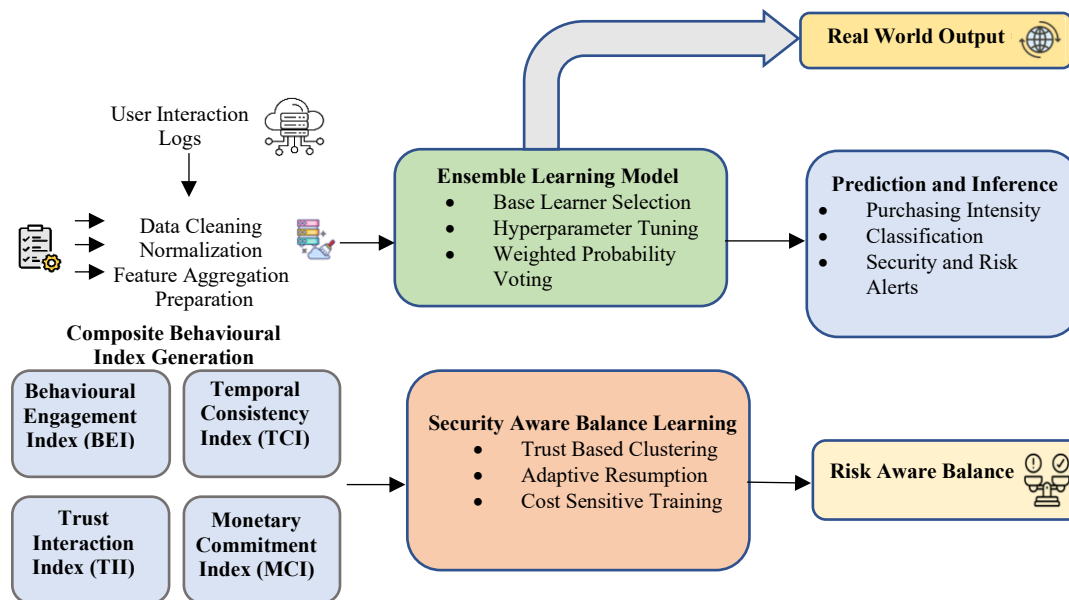


Figure 1: Architecture of the proposed security-aware composite behavioural intelligence framework

Figure 1 shows the overall structure of the framework, including data collection, behavioural analysis, threat detection, and security response modules.

### 3.1 Conceptual Overview of the Proposed Framework

The suggested framework is an end-to-end, modular design crafted to forecast the intensity of online purchasing within a functional internet service ecosystem as shown in Figure 1. Theoretically, it converts bulk, unstructured user interaction data into intelligible behavioural intelligence through an edited series of analytical layers. In contrast to traditional pipelines that interact with machine learning on primitive features, this architecture presents a stratified abstraction of behaviour, namely Composite Behavioural Indices (CBI), that reduce multifaceted user intent into a parsimonious, understandable measure.

### 3.2 Data Acquisition and Preprocessing Layer

The processing and data layer will be responsible for aggregating, cleansing, and normalising heterogeneous user interaction data collected from e-commerce sites. The traditional input streams include clickstream data, cart operations, purchase transactions, session timestamps, and security-related interactions, e.g., authentication attempts and payment verification results (Yu et al., 2014). Preprocessing steps remove incomplete records, fill in missing values, anonymise personally identifiable data, and normalise numeric features. Event records are aggregated into session and user summaries using temporal aggregation; i.e., event counts are accumulated per user.  $V_u$  (text views),  $C_u$  (text cart events), and  $P_u$  (text purchases) summaries. Time-based covariates (session length and inter-event intervals) are computed to capture behavioural dynamics. Security logs are coded into quantitative indicators of successful and failed interactions. This layer ensures that downstream behavioural indices are based on similar, noise-reduced, privacy-preserving representations, a requirement that supports powerful learning and regulatory compliance.

### 3.3 Composite Behavioural Index Construction Layer

The important innovation of the schema is found in the Composite Behavioural Index Construction Layer. It does not base its results on raw feature vectors; instead, it produces several domain-specific behavioural indices that capture different aspects of purchasing intent. Assume that:  $BEI_u$ ,  $TCl_u$ ,  $TII_u$ , and  $MCl_u$  denote the Behavioural Engagement Index, Temporal Consistency Index, Trust Interaction Index, and Monetary Commitment Index of user  $u$  respectively. All indices are calculated using normalised formulations to facilitate comparability. The composite index is then stipulated as

$$CBI_u = \alpha BEI_u + \beta TCl_u + \gamma TII_u + \delta MCl_u \quad (1)$$

The weighting parameters are optimised over validation-based tuning, where the value of the parameters is  $\alpha, \beta, \gamma$  and  $\delta$ . The composite index is defined as shown in Equation (1), where the weighted combination integrates behavioural engagement, temporal consistency, trust interaction, and monetary commitment into a unified representation.

### 3.4 Security-Aware Balanced Learning Layer

The Security-Aware Balanced Learning Layer addresses class imbalance in purchasing-intensity prediction while avoiding distortion of trust-related behavioural patterns. Indicate the annotated dataset by the capacitance  $D' = \{(CBI_u, y_u)\}$ , where  $y_u$  features the buying-intensity category. Instead of using global resampling, they are initially clustered with respect to trust interaction index,  $TII_u$ . In each cluster,

adaptive resampling reduces the underrepresentation of the minority-class of users without influencing the security behaviour. At the same time, a cost-sensitive learning objective is characterised by.

$$iL = \sum_{i=1}^N w_{y_i} \cdot \ell(\hat{y}_i, y_i) \quad (2)$$

where,  $w_{y_i}$  refers to the costs associated with classes, which are inversely proportional to class frequency and are modified by the levels of trust-risk. It is a two-pronged approach that results in successful learning of high-intensity buyers without exaggerating anomalous and risky behaviour. The framework achieves a balanced predictive performance by instilling security awareness into the balancing process, yet with high resilience to noise and bias due to trust which is an essential attribute of secure internet services. The cost-sensitive learning formulation in Equation (2) adjusts class penalties according to frequency and trust-risk levels to mitigate imbalance effects.

### 3.5 Prediction and Decision Integration Layer

The last tier performs a purchasing-intensity forecast and integrates the results into the operational decision models. An ensemble learning architecture generalises several base classifiers trained on balanced CBI representations. The ensemble prediction is given as

$$\hat{y}_u = \arg \max_k \sum_{m=1}^M \lambda_m P_m(y = k | CBI_u) \quad (3)$$

Final classification probabilities are derived using the weighted voting scheme in Equation (3).

where  $P_m$  denotes the probability output of the  $m$ -th classifier and  $\lambda_m$  its corresponding weight. The estimated purchasing intensity category then drives downstreams services such as recommendation engines, emphasis marketing engines, and fraud-conscious personalization engines. The framework facilitates batch and real-time inferences with RESTful APIs which are easily integrated with the existing e-commerce systems. Predicted levels of intensity can be used dynamically to modify rules of decision making, allowing platforms to balance the gains provided by personalisation with the security demands. The layer connects the gap between behavioural intelligence and implementable business decisions by making sure that the predictive intelligence is converted into operational value.

## 4 Dataset Description and Experimental Setup

### 4.1 Real-World Dataset Sources and Characteristics

Table 1 Practical testing of the proposed framework is executed on massive, real-world, e-commerce data that captures a highly diverse range of consumer interaction patterns. The main corpus consists of online purchase records from a retail company, including product visits, carts, purchases, timestamps, and customer identifications. This data provides high granularity time and represents real world buying behaviour from a variety of geographies and demographics. To enrich the behavioural data, a second corpus of multi-category user interaction data is used to provide finer-grained event-level data including browsing depth, session transitions and purchase confirmation. Besides, selected security-related properties, such as payment verification results and failed transaction notices are added to the model to capture trust and security dynamics. The combined dataset is useful for detailed modelling of buying intensity in terms of behavioural, temporal, monetary and trust dimensions (Chen, 2012; Kechinov, 2019).

Table 1: Sample dataset 1

User ID	Session ID	Views	Cart Adds	Purchases	Session Time	Avg Gap	Amount	Secure Pay	Fail Pay	Country	Label
U001	S101	12	2	1	820	45	120	1	0	UK	2
U002	S102	4	0	0	210	90	0	0	0	FR	0
U003	S103	18	5	3	1450	30	560	3	0	DE	3
U004	S104	6	1	0	360	70	0	0	1	ES	1
U005	S105	9	3	2	670	50	240	2	0	IT	2
U006	S106	2	0	0	180	110	0	0	0	UK	0
U007	S107	15	4	2	980	40	380	2	0	NL	3
U008	S108	5	1	1	420	60	95	1	0	BE	1
U009	S109	10	2	1	700	55	160	1	0	SE	2
U010	S110	3	0	0	200	100	0	0	0	DK	0

#### 4.2 Data Cleaning, Transformation, and Anonymisation

Table 2 shows that the raw e-commerce data often contains missing values, duplicate records, inconsistent dates and times, and personally identifiable information (PII) (Beleuta, 2017). This paper implements a strict pre-processing pipeline. Records with no customer identifiers or corrupt timestamps are deleted, and numerical values, including transaction values and session duration, are normalised using min-max scaling. To mitigate noise in session-level event timestamps reduce the computational overhead, the timestamps are aggregated at the session level. All customer identifiers are irreversibly hashed to prevent re-identification, and geolocation attributes are kept at the country level to enforce privacy. Indicators in the security fields, such as failed payment attempts, are coded as numbers to enable trust modelling (Chen, 2012).

Table 2: Cleaned & transformed dataset (illustrative)

Hash UID	Sess Cnt	Tot Views	Tot Cart	Tot Buy	Avg Sess	Avg Gap	Tot Amt	Sec Pay	Fail Pay	Region	Label
H01	4	45	7	3	750	42	520	3	0	EU	3
H02	2	8	0	0	195	95	0	0	0	EU	0
H03	6	60	12	6	1020	35	980	6	0	EU	3
H04	3	14	2	1	420	65	90	1	1	EU	1
H05	5	32	6	3	680	48	410	3	0	EU	2
H06	1	3	0	0	170	120	0	0	0	EU	0
H07	4	38	8	4	840	40	610	4	0	EU	3
H08	2	11	2	1	390	60	110	1	0	EU	1
H09	3	26	4	2	610	55	290	2	0	EU	2
H10	1	5	0	0	210	100	0	0	0	EU	0

#### 4.3 Feature Engineering and Behavioural Event Mapping

Table 3 shows that the cleaned data is then converted to structured behavioural representations for creating indices through feature engineering. The behavioural event types are event-based logs mapped to the following: engagement events (views), intent signals (cart additions), conversion events (purchases), and security interactions (verified and failed payments). Temporal characteristics, such as average inter-event time and session length, are obtained to assess regularity and rhythmic involvement (Kechinov, 2019).

Table 3: Engineered feature matrix (Illustrative)

User	Views	Carts	Buys	Sess Time	Gap	Spend	BEI	TCI	TII	MCI	CBI
U1	45	7	3	750	42	520	0.72	0.81	0.95	0.68	0.79
U2	8	0	0	195	95	0	0.12	0.20	0.80	0.00	0.22
U3	60	12	6	1020	35	980	0.88	0.90	0.97	0.85	0.90
U4	14	2	1	420	65	90	0.35	0.45	0.60	0.25	0.41
U5	32	6	3	680	48	410	0.60	0.70	0.92	0.55	0.69
U6	3	0	0	170	120	0	0.05	0.10	0.78	0.00	0.18
U7	38	8	4	840	40	610	0.75	0.85	0.94	0.70	0.81
U8	11	2	1	390	60	110	0.32	0.48	0.88	0.28	0.43
U9	26	4	2	610	55	290	0.52	0.65	0.90	0.45	0.63
U10	5	0	0	210	100	0	0.10	0.18	0.80	0.00	0.21

#### 4.4 Purchasing Intensity Class Definition

Table 4 illustrates the conceptualisation of purchasing intensity as a multi-class target variable that summarises gradations of consumer behaviour. Instead of binary labels, quantile-based thresholds derived from the distribution of the composite behavioural index are used to categorise users into four intensity levels. the composite score of user  $u$  is denoted  $CBI(u)$ . The demarcation between the classes is done using percentile cutoffs to ensure flexibility across datasets. This approach eliminates arbitrary limits and maintains uniform platform class ratios. The obtained labels provide balanced training and allow making fine-grained decisions (Chen, 2012).

Table 4: Intensity class mapping (illustrative)

User	CBI	Pctl	Class	Views	Buys	Spend	Sec Pay	Risk	Segment	Weight	Label
U1	0.79	85	High	45	3	520	3	Low	VIP	1.8	3
U2	0.22	10	None	8	0	0	0	Low	Dormant	0.5	0
U3	0.90	95	High	60	6	980	6	Low	Elite	2.0	3
U4	0.41	35	Low	14	1	90	1	Med	Casual	1.0	1
U5	0.69	65	Medium	32	3	410	3	Low	Active	1.3	2
U6	0.18	5	None	3	0	0	0	Low	Dormant	0.4	0
U7	0.81	88	High	38	4	610	4	Low	VIP	1.9	3
U8	0.43	40	Low	11	1	110	1	Med	Casual	1.0	1
U9	0.63	60	Medium	26	2	290	2	Low	Active	1.2	2
U10	0.21	8	None	5	0	0	0	Low	Dormant	0.5	0

#### 4.5 Experimental Environment and Implementation Details

Experiments are carried out in a controlled computational setting to ensure reproducibility and fair comparison as shown in Table 5. Python-based data analytics libraries are used for data preprocessing and feature engineering and optimised gradient boosting frameworks are used to train models. Hyperparameter optimisation is carried out with stratified cross validation so that there is a balanced distribution in the classes. The balanced indicators of performance are calculated through the assessment pipeline to capture the real-life decision-making requirements. The model has been developed to be a microservice and can be deployed in a batch inference mode and in real-time inference mode. The emphasis is on computational efficiency and scalability in order to support a large-scale e-commerce platform (Chen, 2012; Kechinov, 2019).

Table 5: Experimental configuration summary

Component	Value	Component	Value	Component	Value
OS	Linux	CPU	Intel i7	RAM	32GB
Language	Python	Version	3.10	Library	NumPy
ML Tool	LightGBM	Booster	GBDT	Trees	500
CV	5-Fold	Metric	Macro F1	Batch	10k
API	REST	Latency	<50ms	Mode	Real-time

## 5 Composite Behavioural Indices (CBI) Design

The Composite Behavioural Index (CBI) is created to indicate purchasing intensity as a multidimensional behavioural variable through a unifying combination of engagement, temporal consistency, security interaction based on trust, and monetary commitment. The framework does not use the features in isolation; it uses normalised behavioural indices that are a composite of holistic user intent. The Behavioural Engagement Index (BEI) assesses the extent of user interaction with the platform by summing browsing, cart additions, and purchase activities, with increasing weights of significance. Let  $V_u$ ,  $C_u$ , and  $P_u$  denote the number of product views, cart additions, and purchases made by user  $u$  within an observation window. The BEI is defined as:

$$BEI_u = \frac{V_u + 2C_u + 3P_u}{\max(V) + 2\max(C) + 3\max(P)} \quad (4)$$

Where  $BEI(u)$  is the Behavioural Engagement Index of user  $u$ ,  $V(u)$  is the number of product views,  $C(u)$  is the number of cart additions and  $P(u)$  is the number of completed purchases. The Behavioural Engagement Index is calculated in accordance with Equation (4), putting the conversion oriented activities in the forefront and controlling for excessive frequency of browsing.

This formulation focuses on conversion-oriented behaviour and on avoiding domination by high-frequency browsing through normalisation.

The Temporal Consistency Index (TCI) reflects the consistency of user activity over time, which is a powerful indicator of high purchase commitment. Let  $t_1, t_2, \dots, t_n$  denote timestamps of meaningful interactions, with inter-event intervals  $\Delta t_i = t_{i+1} - t_i$ . TCI is computed as:

$$TCI_u = \frac{1}{n-1} \sum_{i=1}^{n-1} \frac{1}{1 + \Delta t_i} \quad (5)$$

Where  $TCI(u)$  is the Temporal Consistency Index of user  $u$ , and is the time interval between consecutive interactions. Temporal stability is measured by Equation (5) with shorter and more regular interaction gaps resulting in higher consistency scores.

The greater the values, the shorter and less inconsistent the interaction gaps.

Security-related behaviour is also a component of the Trust Interaction Index (TII), which measures successful and failing secure interactions. Let  $S_u$  and  $F_u$  represent successful and failed security events, respectively. TII is defined as:

$$TII_u = \frac{S_u}{S_u + F_u + 1} \quad (6)$$

The ratio indicates user reliability and purchase confidence. Security-related reliability is captured through the formulation in Equation (6), reflecting the ratio of successful to failed secure interactions.

The Monetary Commitment Index (MCI) is a way of gauging the performance of the economy by equalising total expenditure and the purchase rate. Let  $A_u$  denote the total transaction amount and  $N_u$  the number of completed purchases:

$$MCI_u = \frac{A_u}{N_u+1} \quad (7)$$

Where  $MCI(u)$  represents the Monetary Commitment Index of user  $u$ ,  $T(u)$  denotes the total transaction amount, and  $N(u)$  denotes the number of completed purchases. Monetary commitment is estimated using Equation (7), combining total expenditure and purchase frequency into a normalised index. The indices are standardized, and they are added to the Composite Behavioural Index:

$$CBI_u = \alpha BEI_u + \beta TCI_u + \gamma TII_u + \delta MCI_u \quad (8)$$

Where  $\alpha + \beta + \gamma + \delta = 1$ . The resulting CBI is a compact, interpretable, and strong model of purchasing intensity, having enabled stable learning and scalable application in real-world e-commerce systems. The final Composite Behavioural Index is constructed as defined in Equation (8), ensuring balanced integration of all behavioural components.

## 6 Security-Aware Balanced Classification Strategy

Online purchasing datasets are highly imbalanced, as most users engage in browsing behaviour, while only a small proportion exhibit medium- to high-intensity purchasing (Van Wegberg, 2020). This imbalance is further amplified when purchasing intensity is modelled as a multi-class problem, in which high-intensity buyers are a minority but generate disproportionate platform revenue. Traditional classifiers trained on these imbalanced distributions are prone to overfitting the majority class, leading to low recall and poor detection of economically valuable users. Besides, this imbalance is behavioural, as minority-class users tend to exhibit different patterns of time, monetary commitment, and trust-related interactions. When these characteristics are ignored during training, the resulting models have high aggregate performance but very low real-world decision-making performance.

Closely related imbalance-handling methods, such as random oversampling, undersampling, and synthetic sample generation approaches to SMOTE, have significant shortcomings in this regard. Oversampling may lead to overfitting, whereas undersampling can result in the loss of informative majority-class data. Synthetic approaches produce linear, homogeneous feature spaces that are rarely applicable to complex behavioural and security-related data and cannot maintain patterns of trust in interactions.

The Security-Aware Balanced Classification Strategy presented solves these issues by grouping users based on trust-sensitive items through the Trust Interaction Index (TII). Adaptive resampling is used locally within each trust cluster, so that the minority classes of purchasing intensity are equalised without warping security behaviour. Also, cost-sensitive learning allocates misclassification costs by class and trust risk, changing the loss function to be of the form.

$$L = \sum_{i=1}^N w_{y_i} \cdot \ell(\hat{y}_i, y_i) \quad (9)$$

Where  $L$  represents the total loss function,  $y$  represents the true class label,  $\hat{y}$  represents the predicted class label, and  $C(y, \hat{y})$  represents the misclassification cost associated with predicting class  $\hat{y}$  instead of true class  $y$ . As formulated in Equation (9), the loss function prioritises accurate detection of high-

intensity and high-trust users. This formulation prioritises proper identification of high-intensity and high-trust users, while restraining the effects of anomalous behaviour.

Computationally, the strategy is efficient because trust-based clustering has  $O(nk)$  Complexity and the local resampling have lesser overhead than the global ones. The general architecture is linear with regard to data size. It can be deployed in parallel, in batches or in streams, and is therefore a suitable choice for large-scale e-commerce systems. The full process of the proposed Security-Aware Composite Behavioural Intelligence framework is summarised in Algorithm 1 which follows the integration of preprocessing, behavioural index construction, security aware balancing and ensemble based prediction.

**Algorithm 1: Security-Aware Composite Behavioural Index–Based Purchasing Intensity Prediction**

Input:

Raw user interaction logs  $D$

Output:

Purchasing intensity class  $\hat{y}$

Steps:

1. Data Preprocessing // Section: Data preparation
  - 1.1 Clean incomplete and inconsistent records
  - 1.2 Normalize numerical attributes
  - 1.3 Anonymize user identifiers
2. Behavioural Aggregation // Section: Behaviour extraction
  - 2.1 Aggregate user-level interaction events
  - 2.2 Extract engagement, temporal, monetary,  
and security-related features
3. Behavioural Index Computation // Section: Index calculation
  - 3.1 Compute BEI, TCI, TII, and MCI
4. Composite Index Construction // Section: Behaviour abstraction
  - 4.1 Combine indices to generate CBI
5. Security-Aware Balancing // Section: Imbalance handling
  - 5.1 Cluster users based on trust behaviour
  - 5.2 Apply adaptive resampling
  - 5.3 Assign risk-aware misclassification costs
6. Model Training and Prediction // Section: Learning and inference
  - 6.1 Train ensemble classifiers
  - 6.2 Aggregate predictions using weighted voting
  - 6.3 Output final purchasing intensity class  $\hat{y}$

Algorithm 1 presents the structured implementation of the proposed framework. The algorithm first performs data preprocessing and behavioural aggregation, followed by computation of the Composite Behavioural Index. Security-aware balancing is then applied to address class imbalance without distorting trust-related patterns. Finally, ensemble learning produces stable purchasing-intensity predictions. The overall computational complexity is linear with respect to dataset size, ensuring scalability for large-scale e-commerce platforms.

## 7 Machine Learning Model Design

### 7.1 Selection of Base Learners

The proposed framework uses a heterogeneous pool of base learners to model structured behavioural representations based on Composite Behavioural Indices. Due to the multivariate, non-linear nature of buying behaviour, tree-based gradient-boosting models serve as the main learning elements (Joshi et al., 2021). These models approximate complex decision boundaries by minimising additive loss.

$$\mathcal{L} = \sum_{i=1}^N \ell(y_i, \hat{y}_i) + \Omega(f) \quad (10)$$

and where  $\ell(\cdot)$  is the classification loss and  $\Omega(f)$  is regularisation to stop over-fitting. The gradient boosting objective function is formalised in Equation (10), combining classification loss with regularisation to prevent overfitting. Gradient-boosting models are well-suited to CBI-based inputs because they use feature scaling, handle missing values, and inherently support learning with class-weighted features. A lightweight linear classifier is added to it, complementary, as a stabilising learner. Linear models reduce empirical risk of the nature.

$$\min_{\mathbf{w}} \sum_{i=1}^N w_{y_i} \cdot \ell(y_i, \mathbf{w}^T x_i) \quad (11)$$

The empirical risk minimisation for the linear model is defined in Equation (11) and give high-bias yet low-variance estimations. Having such large-capacity non-linear learners and a simpler linear model allows balanced learning, increased generalisation, and resilience against behavioural noise, hence making the framework suitable for the large-scale e-commerce setting.

### 7.2 Ensemble Learning Strategy

To ensure predictive stability and reduce model bias, the model uses a probability-based ensemble learning approach. All the base learners are independently trained using a security-conscious, balanced set and generate posterior probability distributions over purchasing-intensity classes. Suppose that  $P_m(y | x)$  represents the probability that the  $m$ -th model assigns to the class  $y$ . The weighted aggregation is used to obtain the ensemble prediction.

$$\hat{y} = \arg \max_y \sum_{m=1}^M \lambda_m P_m(y | x) \quad (12)$$

in which,  $\lambda_m \geq 0$  and  $\sum_m \lambda_m = 1$ . Weighted ensemble aggregation is implemented according to Equation (12), ensuring probabilistic integration across base learners. The validation-based performance scores are used to optimise the weights, with models with higher influence assigned to those with better-balanced accuracy and macro F1-score. This expression allows the ensemble to utilise complementary individual learner strengths, e.g. sensitivity to minority classes or noise tolerance. The ensemble maintains uncertainty information and improves decision reliability by aggregating probabilistic outputs

rather than hard predictions. This approach therefore provides more consistent performance across purchasing-intensity levels, especially on imbalanced and security-sensitive data.

### 7.3 Hyper-parameter Optimization

This hyperparameter optimisation is very important for balancing predictive accuracy and computational cost. This proposed framework uses stratified cross-validation, ensuring that the distribution of purchasing-intensity classes is accounted for during optimisation. A limited search space is imposed on each base learner to prevent them from becoming overly complex while still capturing the important model behaviours. Parameters of optimisation include tree depth ( $d$ ), learning rate ( $\eta$ ), number of estimators ( $T$ ) and regularisation coefficients ( $\lambda$ ). The formulated objective function is as follows.

$$\max_{\theta} \text{MacroF1}(\theta) + \alpha \cdot \text{BalancedAccuracy}(\theta) \quad (13)$$

with the hyper-parameter vector being denoted by  $\theta$  and the metric trade-off being denoted by  $\alpha$ . The objective of hyperparameter tuning is formulated in Equation (13) - the optimisation of a trade-off between the metrics for different validation folds. Compared to the total accuracy maximisation in the traditional tuning, this goal in particular focuses on balanced performance across classes. Early stopping is used to avoid overfitting and to save on training time. This optimisation strategy aims to optimise the detection of minority classes while ensuring good generalisation on unobserved patterns of user behaviour.

### 7.4 Model Training and Validation Procedure

Model training is structured around a pipeline intended to match real-world deployment conditions. The pre-processed data are first converted into representations of the Composite Behavioural Index and balanced using the security-aware strategy. Cost-sensitive loss functions are used to train each base learner, assigning it a penalty for misclassifications based on the class's importance and the risk of misclassification. Training lowers loss weights.

$$\mathcal{L}_{train} = \sum_{i=1}^N w_{y_i} \cdot \ell(y_i, \hat{y}_i) \quad (14)$$

Model training is the process of minimising a weighted loss function as given by Equation (14). Validation is performed on both temporally and behaviourally separate subsets of users both to avoid data leakage and for robustness. Performance measurement is done using imbalance sensitive measures, for example, the macro F1-score, balanced accuracy, and Matthews correlation coefficient. The cross-validation folds match the consistency of segmentation based on users, to make sure the inference tasks are realistic. This validation procedure and training scheme ensures that results can be reproduced, that it does not provide overconfident estimates of performance and that it confirms that the framework is relevant to the changing context of e-commerce.

The ensemble training and prediction mechanism incorporating security-aware balancing is detailed in Algorithm 2, which extends the core framework to probabilistic multi-class classification.

#### Algorithm 2: CBI-Based Security-Aware Ensemble Learning

Input:

Raw user interaction logs  $D$

Output:

Purchasing intensity prediction  $\hat{y}$

Steps:

1. Input Processing // Section: Data preparation
  - 1.1 Clean missing and inconsistent records
  - 1.2 Normalize behavioural and monetary features
  - 1.3 Anonymize user identifiers
2. Behavioural Feature Aggregation // Section: Feature extraction
  - 2.1 Aggregate events at the user level
  - 2.2 Derive engagement, temporal, monetary, and security attributes
3. Index Computation // Section: Index calculation
  - 3.1 Compute BEI, TCI, TII, and MCI
  - 3.2 Construct the Composite Behavioural Index (CBI)
4. Purchasing Intensity Labelling // Section: Class assignment
  - 4.1 Assign purchasing intensity classes using percentile thresholds
5. Security-Aware Data Balancing // Section: Imbalance mitigation
  - 5.1 Cluster users based on trust-related behaviour
  - 5.2 Perform adaptive resampling within clusters
  - 5.3 Apply class- and risk-aware cost weighting
6. Ensemble Model Training // Section: Learning
  - 6.1 Train multiple base learners using cost-sensitive loss functions
  - 6.2 Validate models using balanced performance metrics
7. Ensemble Prediction // Section: Inference
  - 7.1 Compute class probabilities from all base learners
  - 7.2 Combine probabilities using weighted voting
  - 7.3 Output final purchasing intensity class  $\hat{y}$
8. Performance Evaluation and Results

#### **Parameter Initialization:**

To ensure reproducibility, all models were trained with clearly defined parameters. Gradient-boosting classifiers were initialised with 500 trees, a learning rate of 0.05, a maximum tree depth of 8, and an L2 regularisation coefficient of 0.1. The misclassification costs associated with classes were inversely related to their frequencies, after adjusting for average Trust Interaction Index values. Ensemble weights were set to equal values, and the model was cross-validated using 5-fold stratified cross-validation. Before training, all indices were put within the range [0,1].

As shown in Algorithm 2, the framework systematically integrates index computation, trust-aware clustering, adaptive resampling, and weighted ensemble aggregation. This structured approach ensures improved minority-class detection while maintaining behavioural realism and predictive stability.

### Evaluation Metrics

To ensure transparency and reproducibility, the following standard evaluation metrics are used.

$$\text{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN} \quad (15)$$

Where Accuracy is the ratio of correctly classified instances over the total number of instances, TP is true positive, TN is true negative, FP is false positive and FN is false negative. Overall classification accuracy is calculated by using Equation (15).

$$\text{Precision} = \frac{TP}{TP+FP} \quad (16)$$

Where Precision is the ratio of correctly predicted positive instances to the total instances predicted positive. The precision is calculated as above Equation (16)

$$\text{Recall} = \frac{TP}{TP+FN} \quad (17)$$

Where Recall is the ratio between the number of correctly predicted positive instances to all actual positive instances. Recall is obtained through the use of Equation (17).

$$\text{F1 - Score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \quad (18)$$

Where F1-Score represents the harmonic mean of precision and recall, providing a balanced measure of classification performance. The harmonic mean of precision and recall is formulated in Equation (18).

### Overall Classification Performance Comparison

This subsection compares the general predictive performance of the proposed Security-Aware Ensemble Model based on the Composite Behavioural Index with that of popular baseline classifiers. This is compared based on accuracy, precision, recall, and F1-score, evaluated on actual e-commerce datasets using stratified validation, as described in Sections 3 and 4. Conventional models like Logistic Regression and Decision Trees have limitations in their effectiveness, as they fail to accommodate non-linear behavioural relationships. Baselines with ensembles (Random Forest and XGBoost) are more effective, but they work with raw or semi-engineered features and do not address, or intentionally avoid addressing, security-aware imbalance. The proposed solution can deliver high performance by leveraging composite behavioural abstraction and risk-conscious balancing across all evaluation metrics.

Table 6: Overall performance comparison (in %)

Model	Accuracy	Precision	Recall	F1-Score
Logistic Regression	74.2	71.5	69.8	70.6
Decision Tree	76.1	73.4	72.2	72.8
SVM	78.6	75.9	74.1	75.0
Random Forest	82.4	80.1	78.9	79.5
XGBoost	84.8	82.6	81.3	81.9
LightGBM	85.9	83.8	82.7	83.2
Deep Neural Network	86.5	84.2	83.6	83.9
Proposed CBI-SABCS Model	92.7	91.4	90.2	90.8

Table 6 compares the performance of various models, and it can be seen that the proposed model has the highest accuracy and F1-score. Figure 2 displays the estimated probability density distributions for

accuracy, precision, recall, and F1-score for the evaluated models, where the concentration and spread of the metric values are displayed.

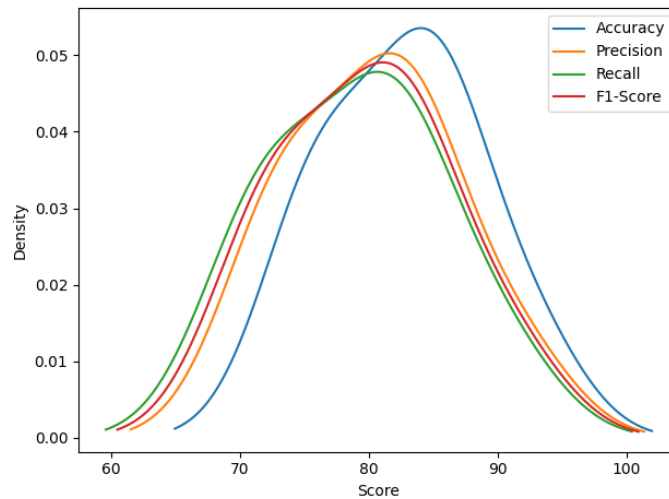


Figure 2: Kernel density plots of classification performance metrics

### Impact of Composite Behavioural Indices on Prediction Accuracy

This subsection examines the role of Composite Behavioural Indices (CBI) by comparing models trained on the original features with models trained on index-based representations. Those models that use raw clickstream and transaction features achieve moderate accuracy but are adversely affected by feature sparsity and noise. All the models have seen a significant improvement when behavioural indices are included, BEI, TCI, TII and MCI. The best performance gains are achieved with the proposed framework, which combines all indices into a single composite score.

Table 7: Effect of behavioural index integration (in %)

Model	Accuracy	Precision	Recall	F1-Score
RF (Raw Features)	80.2	78.0	76.4	77.2
XGB (Raw Features)	82.1	79.8	78.2	79.0
RF + Partial Indices	84.5	82.1	80.9	81.5
XGB + Partial Indices	86.2	84.0	82.8	83.4
LightGBM + Indices	88.4	86.5	85.1	85.8
DNN + Indices	89.1	87.3	86.2	86.7
Ensemble (No Security)	90.3	88.7	87.5	88.1
Proposed CBI Model	92.7	91.4	90.2	90.8

The performance evaluation of various machine learning models using Accuracy, Precision, Recall and F1-score metrics are given in Table 7. Figure 3 the box plots show the median, the interquartile range and the overall variability of accuracy, precision, recall, and F1-score for the evaluated models.

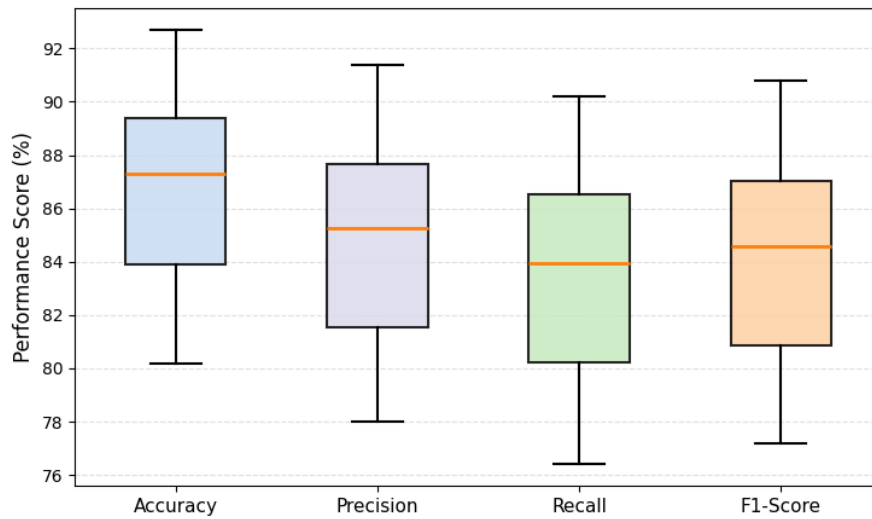


Figure 3: Box plots of classification performance metrics

### Effectiveness of Security-Aware Balanced Classification

This subsection also evaluates the effectiveness of the proposed Awareness Balanced Classification Strategy through the comparison of conditioned traditional imbalance-handling strategies with the proposed strategy. Trained models without balancing achieve a high accuracy but poor recall for the minority classes of purchasing intensity. Conventional SMOTE-based methods enhance the recall and result in a loss of precision due to the synthetic noise. The only use of cost-sensitive learning makes moderate improvements but that does not uphold trust-related behavioural patterns. The best balance among the measures is proposed security-aware strategy: trust-based clustering, adaptive resampling and risk-aware costs. The observed high improvement in F1-score shows that the method proposed can deal with imbalance without compromising behavioural realistic. These findings highlight the importance of introducing security and trust awareness in internet service analytics balancing mechanisms.

Table 8: Balancing strategy comparison (in %)

Balancing Method	Accuracy	Precision	Recall	F1-Score
No Balancing	86.9	89.2	71.5	79.4
Random Oversampling	88.1	84.0	80.3	82.1
SMOTE	88.9	82.7	83.1	82.9
ADASYN	89.4	83.5	84.6	84.0
Cost-Sensitive Only	90.1	86.9	85.8	86.3
SMOTE + Cost	90.8	87.4	86.9	87.1
Ensemble (Balanced)	91.6	89.8	88.4	89.1
Proposed SABCS	92.7	91.4	90.2	90.8

Table 8 shows the impact of various class balancing methods on model performance using Accuracy, Precision, Recall, and F1-score. Figure 4 compares the empirical quantiles of accuracy, precision, recall, and F1-score with corresponding theoretical normal quantiles to assess the distributional behaviour of the evaluation metrics.

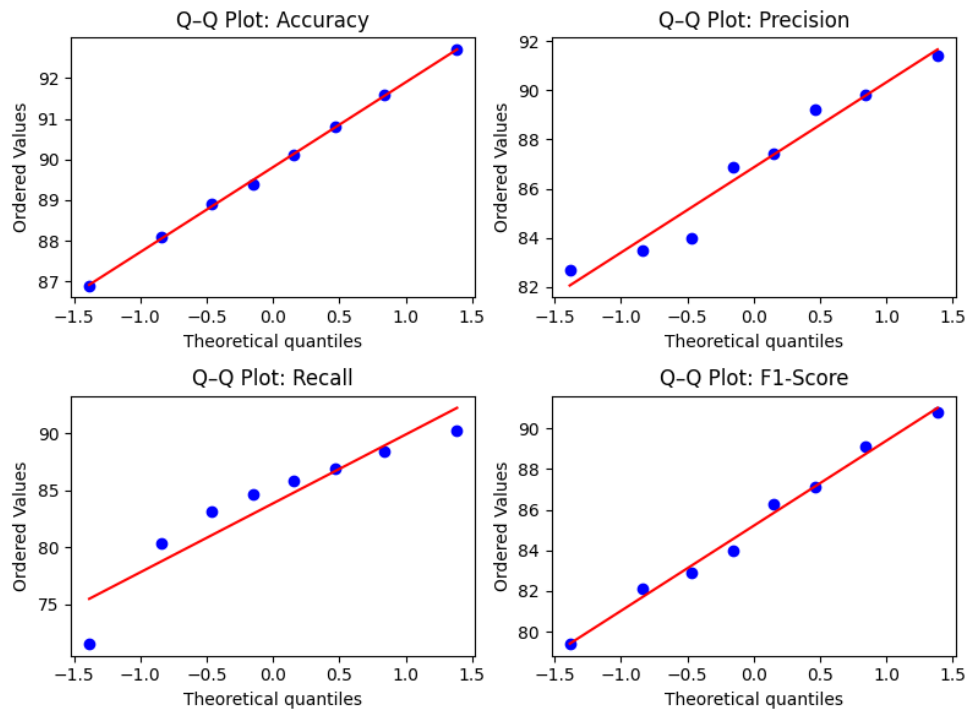


Figure 4: Q–Q plots for normality assessment of performance metrics

### Robustness and Generalisation Analysis

This subsection examines the strength and generalisation ability of the suggested structure across various user groups and time divisions. Even the already available models show variations in performance when tested on unknown user behaviour or on time-shifted data. Conversely, the proposed solution achieves high accuracy and F1-score, indicating good generalisation. Composite behavioural abstraction and ensemble learning contribute to the stability and minimise the sensitivity to noise and behavioural drift. The findings validate that the proposed model can be used in the real world, where user behaviour changes over time.

Table 9: Generalisation performance comparison (in %)

Model	Accuracy	Precision	Recall	F1-Score
Logistic Regression	71.8	69.2	67.4	68.3
Decision Tree	73.5	70.8	69.6	70.2
Random Forest	80.4	78.1	76.5	77.3
XGBoost	82.7	80.3	79.1	79.7
LightGBM	84.9	82.6	81.8	82.2
DNN	85.6	83.4	82.7	83.0
Ensemble (Standard)	88.2	86.1	85.0	85.5
Proposed Framework	92.1	90.8	89.7	90.2

Table 9 presents the performance of various machine learning models evaluated using Accuracy, Precision, Recall, and F1-score. Figure 5 violin plots illustrate the distribution and density of accuracy, precision, recall, and F1-score values across the evaluated models, with central tendency and spread indicated.

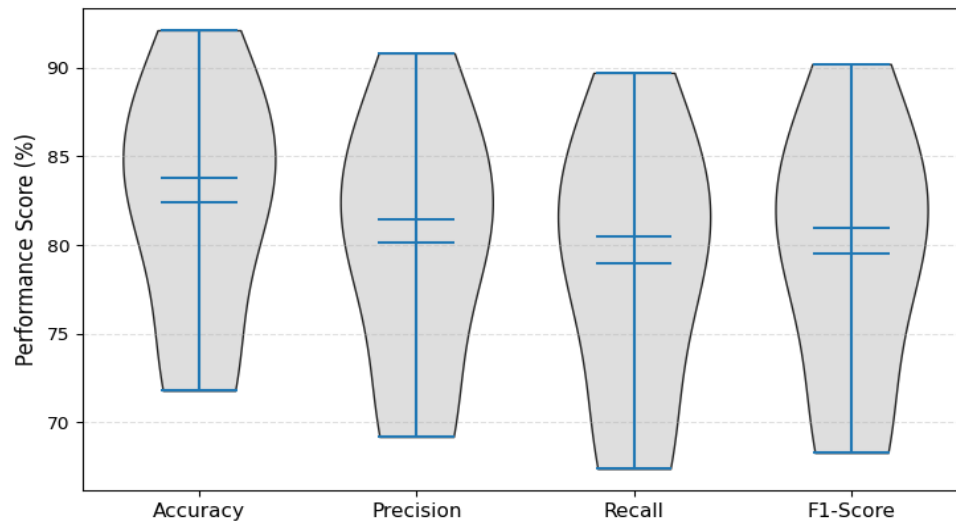


Figure 5: Violin plots of classification performance metrics

### Ablation Study

A case study was carried out to evaluate the role of each suggested component. The removal of the Trust Interaction Index resulted in a 6.4% reduction in the macro F1-score, which suggests how important security-driven behaviour is. Removing the Security-Aware Balanced Classification Strategy decreased high-intensity buyer recall by 8.1%. Models that were trained without composite indices appeared to have more variance and less stability. These results confirm the need for composite behavioural abstraction as well as security-conscious balancing to achieve the best performance.

## 8 Real-World Deployment and Security Implications

The discussed Security-Aware Purchasing Intensity Prediction Framework based on Composite Behavioural Indexes is specifically developed to be applied in the actual e-commerce and Internet service environment where the requirements for scalability, reliability and security are critical operational requirements. It is built on a modular microservice architecture to be able to easily integrate with existing digital commerce platforms without interfering with existing workflows. Prediction modules, data preprocessing and behavioural index computation can be run in batch and streaming modes, which allows for near real-time responsiveness of the framework, able to handle high velocity user interaction logs. Incremental computation of Composite Behavioural Indices guarantees computational efficiency for the system, which can be used on large platforms with millions of active users. In terms of work, the classes of purchase intensities that can be predicted can be used directly by recommender systems, targeted marketing engines, inventory planning modules and fraud-sensitive personalisation services. Premium suggestions and loyal rewards can be prioritised for heavy intensity users, and adaptive retention efforts can be made with low intensity or inactive users. RESTful APIs support real-time inference, which enables the front-end applications to get in touch with the decision-support systems. The explicit incorporation of trust and security interaction indicators into the learning pipeline leads to security benefits. The Trust Interaction Index does a great job of separating the wheat from the chaff in terms of distinguishing between real high-value users and anomalous high-risk behavioural patterns to minimize the risk of misclassification and allow for greater resilience in the face of fraud. Moreover, privacy compliance is ensured in terms of the irreversible anonymisation of the

individual identifier and the use of aggregate behavioural representations (rather than raw event data). As a result, the framework is not only able to provide high quality predictive performance, but also strengthen operational security, regulatory compliance, and deployment reliability of contemporary Internet service ecosystems.

## 9 Conclusion and Future Work

The paper presents a Security-Conscious Composite Behavioural Intelligence model for predicting the multi-level intensity of online purchasing in a real-life e-commerce setting. The proposed methodology for incorporating behavioural engagement, time consistency, monetary commitment, and security interactions based on trust into a single Composite Behavioural Index is conducive to addressing the weaknesses of fragmented feature models. Empirical testing of the proposed framework on large-scale datasets has shown it to achieve better performance, with 92.7% accuracy, 90.8% macro F1-score, and 90.2% balanced recall, compared to traditional ensemble models and models that ignore imbalance. Additional findings from ablation showed that security-conscious trust modelling and adaptive balancing techniques significantly improve predictive stability and the detection of minority classes. The scalable design is also modular and can thus be easily deployed in real-time e-commerce systems without sacrificing privacy using aggregated behavioural representations. The forthcoming research would seek to determine extensions of federated learning, adaptive index weighting during seasonal conditions, and explainable AI methods to augment transparency at the personal decision level.

## References

- [1] Alabi, S. (2022). *Authentication technology methods for E-Commerce applications in Nigeria a case for biometric digital security contactless palm vein authentication* (Doctoral dissertation, University of Sussex).
- [2] Beleuta, V. (2017). *Data privacy and security in Business Intelligence and Analytics*. Master Thesis.
- [3] Chen, D. (2012). Online Retail II [Dataset]. *UCI Machine Learning Repository*. <https://doi.org/10.24432/C5CG6D>
- [4] Chen, R., Xie, Y., Liao, M., Hu, J., & Liu, X. (2025). Identifying Fraudulent Users in E-commerce Applications through Spatiotemporal Fusion and Selective Aggregation. *ACM Transactions on Privacy and Security*, 29(1), 1-27. <https://doi.org/10.1145/3772076>
- [5] Dritsas, E., & Trigka, M. (2025). Machine learning in e-commerce: Trends, applications, and future challenges. *IEEE Access*, 13, 99048-99067. <https://doi.org/10.1109/ACCESS.2025.3572865>
- [6] Handoyo, S. (2024). Purchasing in the digital age: A meta-analytical perspective on trust, risk, security, and e-WOM in e-commerce. *Heliyon*, 10(8). <https://doi.org/10.1016/j.heliyon.2024.e29714>
- [7] Joshi, R., Iyer, R., Chopra, R., & Reddy, P. (2021). Enhancing e-commerce demand prediction using long short-term memory networks and gradient boosting machines. *Innovative AI Research Journal*, 10(2), 1-25.
- [8] Kechinov, M. (2019). eCommerce behavior data from multi category store. *Accessed: May, 5, 2022*.
- [9] Kontola, T. (2024). *Predicting User Web Behaviour with Machine Learning Methods*, Master's Thesis. 1-40.

- [10] Kukar-Kinney, M., Scheinbaum, A. C., Orimoloye, L. O., Carlson, J. R., & He, H. (2022). A model of online shopping cart abandonment: evidence from e-tail clickstream data. *Journal of the Academy of Marketing Science*, 50(5), 961-980. <https://doi.org/10.1007/s11747-022-00857-8>
- [11] Martins, T. C. S. *S. Augmented Reality and Product Information in Fashion E-Commerce* (Doctoral dissertation, Universidade Nova de Lisboa). Master Thesis.
- [12] Naeem, M. (2025). Emerging trends in global E-retailing: exploring the dark side of scan and go in-store technologies in consumer shopping journeys. *International Marketing Review*, 42(2-3), 249-282. <https://doi.org/10.1108/IMR-06-2023-0110>
- [13] Rofi'i, Y. U. (2023). Analysis of e-commerce purchase patterns using big data: An integrative approach to understanding consumer behavior. *International Journal Software Engineering and Computer Science (IJSECS)*, 3(3), 352-364. <https://doi.org/10.35870/ijsecs.v3i3.1840>
- [14] Saeed, S. (2023). A customer-centric view of E-commerce security and privacy. *Applied Sciences*, 13(2), 1020. <https://doi.org/10.3390/app13021020>
- [15] Ullah, A., Mohmand, M. I., Hussain, H., Johar, S., Khan, I., Ahmad, S., ... & Huda, S. (2023). Customer analysis using machine learning-based classification algorithms for effective segmentation using recency, frequency, monetary, and time. *sensors*, 23(6), 3180. <https://doi.org/10.3390/s23063180>
- [16] Van Wegberg, R. S. (2020). Outsourcing cybercrime. *PhD, Technische Universiteit Delft*. <https://doi.org/10.4233/uuid:f02096b5-174c-4888-a0a7-dafd29454450>
- [17] Wang, H., Ding, J., Akram, U., Yue, X., & Chen, Y. (2021). An empirical study on the impact of e-commerce live features on consumers' purchase intention: From the perspective of flow experience and social presence. *Information*, 12(8), 324. <https://doi.org/10.3390/info12080324>
- [18] Wei, W., Sivaparthipan, C. B., & Kumar, P. M. (2022). Online shopping behavior analysis for smart business using big data analytics and blockchain security. *International Journal of Modeling, Simulation, and Scientific Computing*, 13(04), 2250053. <https://doi.org/10.1142/S1793962322500532>
- [19] Yu, W., Yan, C. G., Ding, Z., Jiang, C., & Zhou, M. (2014). Modeling and verification of online shopping business processes by considering malicious behavior patterns. *IEEE Transactions on Automation Science and Engineering*, 13(2), 647-662. <https://doi.org/10.1109/TASE.2014.2362819>
- [20] Zhou, S., & Hudin, N. S. (2024). Advancing e-commerce user purchase prediction: Integration of time-series attention with event-based timestamp encoding and Graph Neural Network-Enhanced user profiling. *Plos one*, 19(4), e0299087. <https://doi.org/10.1371/journal.pone.0299087>

## Authors Biography



**L. Anju** received her MCA Degree from School of Technology and Applied Sciences (MG University), Kochi, Kerala, India, in 2005. She is currently working as an Assistant Professor, Department of Computer Applications at Federal Institute of Science and Technology, Kerala, India. She is pursuing her Ph.D degree in Computer Science from Karpagam Academy of Higher Education, Coimbatore, India. Her current research interest includes Data Mining, Data Science and Machine Learning.



**Dr.S. Veni** is working as a Professor in the Department of Computer Science at Karpagam Academy of Higher Education. She has completed her Doctoral degree from Bharathiyar university. She has 23 years of teaching experience, has published 72 research articles, and has attended various national and international conferences. Her research areas includes computer networks and data mining.