

FraudLens: A Hybrid Machine Learning Framework for Fake Discount Detection in E-Commerce Platforms

Amrithavarshini S¹, Anittajoju N², Anushree A³, Sreya P^{4*}, Anitha M⁵

1,2,3,4,5Department of Computer Science and Engineering, Avinashilingam Institute for Home Science and Higher Education for Women, Coimbatore, Tamil Nadu, India

**Corresponding Author: Sreya P, 23uea047@avinuty.ac.in*

Abstract – E-commerce platforms increasingly rely on discount-based promotions to influence consumer purchasing decisions, particularly during major sale events. This reliance has given rise to a widespread and largely undetected form of pricing fraud, in which sellers artificially inflate the Maximum Retail Price (MRP) prior to a sale and subsequently apply a discount, creating the illusion of substantial savings that do not reflect any genuine price reduction. Existing price-tracking tools such as Keepa and CamelCamelCamel provide descriptive historical price data but lack the diagnostic capability to determine whether a discount is genuine. This paper presents FraudLens, a hybrid, data-driven fraud detection framework that combines an unsupervised anomaly detection layer (Isolation Forest) with supervised classification models (Random Forest and XGBoost) to identify manipulated discounts across Amazon and Flipkart product listings. The system incorporates domain-driven feature engineering — including MRP inflation ratio, discount permanence score, pre-sale price spike detection, and seller trust score — together with SHAP-based explainability to translate model predictions into human-readable justifications. The proposed system was trained and validated on a combined dataset of 73,114 records spanning 3,000 unique products, and the XGBoost classifier achieved a weighted F1-score of 96.9%, correctly classifying products as Genuine, Suspicious, or Fake. An interactive Streamlit dashboard was also developed to provide real-time, per-product fraud verdicts supported by visual evidence, enabling consumers to make more informed purchasing decisions.

Key Words: *Fake Discount Detection, E-Commerce Fraud, Anomaly Detection, Isolation Forest, XGBoost, Explainable AI (SHAP)*

1. INTRODUCTION

E-commerce platforms such as Amazon and Flipkart depend heavily on discount-driven marketing to attract consumers, particularly during large-scale sale events such as Big Billion Day and the Great Indian Festival. Discounts are one of the strongest psychological levers in online retail, and sellers are aware that a steep percentage-off figure directly influences purchase intent, regardless of whether the final price reflects genuine value.

This dependence on discount marketing has created an opportunity for manipulation. A common tactic observed across online marketplaces is MRP inflation, where a seller artificially raises the listed Maximum Retail Price shortly before a sale event and then applies a discount against this inflated baseline. The final selling price may be equal to or

even higher than the product's historical market price, yet it is displayed as a "steep discount." This practice misleads consumers into believing they are receiving genuine savings, erodes trust in online marketplaces, and results in inefficient purchase decisions driven by perceived rather than actual value.

Currently, no automated, verdict-generating system exists for ordinary shoppers to check whether a displayed discount is real. Available tools largely function as descriptive price trackers rather than diagnostic fraud detectors — they show historical price movement but do not analyze the intent or pattern behind a price change, and they remain vulnerable to dark-pattern tactics such as strikethrough pricing and pre-sale price spikes.

To address this gap, this paper proposes FraudLens, a hybrid rule-and-machine-learning based system designed to: (i) detect fraudulent discount practices by identifying artificial MRP inflation and deceptive pricing patterns through data-driven analysis; (ii) integrate unsupervised anomaly detection (Isolation Forest) with supervised classification (Random Forest, XGBoost) for high-accuracy fraud prediction; (iii) ensure model transparency using Explainable AI (SHAP), translating predictions into clear, interpretable insights; and (iv) provide real-time decision support through an interactive dashboard that visualizes price history, fraud signals, and prediction confidence.

2. LITERATURE SURVEY

Recent research has approached the problem of e-commerce pricing fraud and dark-pattern detection from several angles, each contributing valuable but partial solutions.

A visual-analytics platform, ECommVis, has been proposed to help e-commerce merchants interpret complex advertising and sales metrics through contextual, user-centric visualization [1]. This work demonstrates that visual simplification is critical to stakeholder decision-making on e-commerce platforms, supporting the case for a dashboard-driven fraud verdict system rather than raw data dumps.

Separately, a large-scale text-based dataset of dark-pattern instances has been used to benchmark automated classification models — including BERT, RoBERTa, ALBERT, and XLNet — for detecting manipulative interface language, with the best model reaching 97.5% accuracy [2]. This validates the shift from manual pattern tracking toward automated, model-based classification of deceptive design and pricing behavior, a direction FraudLens extends

specifically to price and discount manipulation rather than interface text.

At industrial scale, a combined unsupervised-and-supervised anomaly detection system has been deployed in a large e-commerce pricing pipeline to flag mis-priced items in both batch and real-time streaming settings [3], demonstrating that anomaly-scoring layers can operate reliably ahead of any rule-based review. This principle — anomaly detection as an upstream signal rather than a final verdict — forms the basis of the unsupervised anomaly-scoring layer used in FraudLens.

Beyond machine learning literature, widely used commercial tools such as Keepa and CamelCamelCamel remain purely descriptive: they display historical price charts but stop short of issuing a fraud verdict, leaving interpretation entirely to the user. Prior large-scale audits of shopping websites have further shown that manipulative interface and pricing patterns — including inflated strikethrough prices — are common and systematically exploited across thousands of online storefronts [4]. Static, rule-only trackers are documented to struggle at distinguishing a natural seasonal price trend from a deliberate pre-sale spike engineered purely to fabricate a discount.

Collectively, this body of work indicates three unresolved gaps that FraudLens is designed to close: (a) the absence of an explicit “deception logic” layer that evaluates intent rather than just historical value, (b) vulnerability to dark-pattern tactics that rely on manipulable website metadata, and (c) reliance on static rule sets that cannot adapt to gradual, camouflaged price inflation.

3. PROPOSED SYSTEM

A. System Architecture

FraudLens is built around a Rule-First Hybrid Architecture, in which a rule-based decision engine acts as the governing logic layer, supported by machine learning models that supply the underlying evidence for each decision. The pipeline enforces a zero-trust detection approach: every discount claim is validated against a 90-day historical price baseline rather than being trusted at face value. A 7-day sensitivity guard specifically targets pre-sale price inflation, while a discount integrity filter flags any deviation greater than 2% between the displayed discount and the actual computed discount.

Machine learning is layered on top of this rule engine rather than replacing it. An Isolation Forest model performs unsupervised anomaly scoring to catch “price creeping” — slow, incremental price hikes designed to evade simple threshold rules — and this anomaly score is then fed as an additional input feature into two supervised classifiers, Random Forest and XGBoost, which produce the final Genuine / Suspicious / Fake verdict.

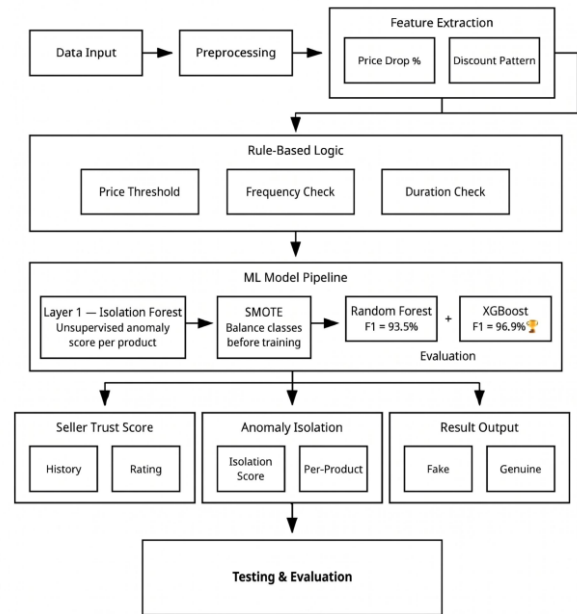


Fig-1: System architecture and modular flow pipeline of FraudLens

As illustrated in Fig-1, the pipeline begins with raw data ingestion, where product listings from Amazon and Flipkart are read into the system as structured Data Input. This raw data first passes through the Preprocessing stage, where missing values, price outliers, and duplicate listings are resolved before any fraud signal is computed. The cleaned data then enters the Feature Extraction block, which independently derives two core signal families: the Price Drop % (the magnitude and direction of change relative to the historical baseline) and the Discount Pattern (the shape and frequency of price changes over the 90-day window). These extracted features feed directly into the Rule-Based Logic block, which performs three lightweight, deterministic checks — a Price Threshold check that flags MRP inflation beyond a fixed tolerance, a Frequency Check that detects how often a discount is applied and withdrawn, and a Duration Check that measures how long a “discounted” price is actually sustained, distinguishing a genuine markdown from one that reverts within days.

Only after clearing this rule layer does data proceed to the ML Model Pipeline, where an Isolation Forest first assigns each product an unsupervised anomaly score, SMOTE rebalances the training data to counter class imbalance, and the balanced feature set is passed in parallel to Random Forest and XGBoost classifiers, whose predictions are combined at the Evaluation stage. The output of this pipeline branches into three parallel modules: a Seller Trust Score module that aggregates seller History and Rating, an Anomaly Isolation module that surfaces the per-product Isolation Score, and a Result Output module that renders the final Fake or Genuine verdict. All three branches converge at a common Testing & Evaluation stage, ensuring that every verdict issued by FraudLens is cross-validated against held-out data before being surfaced to the user. This layered design — rules first, anomaly detection second, supervised classification third — allows the system to remain both fast, since most genuine

listings are cleared by the rule layer alone, and robust, since borderline cases are escalated to the full ML pipeline rather than being misclassified outright.

B. Module-Wise Description

Module 1 – Data Collection and Preprocessing: Amazon and Flipkart datasets were ingested using Pandas, unified into a common schema, and reconciled for product ID conflicts to ensure cross-platform consistency. A validation pipeline enforced price constraints (₹1–₹10,00,000), capped the selling price at 105% of MRP, removed duplicates, and applied time-series-based imputation for missing values. To evaluate detection performance under realistic conditions, controlled discount manipulation was synthetically injected into ~8% of records, creating a deliberate mismatch between displayed and actual discount. A discount validation engine then computed the true discount and flagged any deviation greater than 2% as anomalous. Foundational features — including the 90-day baseline price, MRP inflation ratio, discount permanence, price volatility, and pre-sale spike indicators — were engineered at this stage and persisted into a SQLite database (price_monitor.db).

Module 2 – Feature Engineering and Labeling: Advanced fraud signals were derived, including a round-discount flag, charm-pricing detection (e.g., ₹99-ending prices), listing-age anomaly, seller trust score, and discount gap. These were combined into a composite fraud score (0–10) to strengthen class separability. Exploratory data analysis using Matplotlib and Seaborn (correlation heatmaps, fraud distribution plots, pricing trend charts) validated feature relevance, while cross-platform analysis surfaced abnormal MRP/price gaps between Amazon and Flipkart listings of comparable products. The dataset was then aggregated to the product level, label-encoded, and scaled using Min-Max normalization, and split into Train (70%) / Validation (15%) / Test (15%) sets with balanced fraud-class representation.

Module 3 – Model Building and Detection: A two-layer hybrid architecture was implemented: Isolation Forest for unsupervised anomaly scoring (contamination ≈ 20%), followed by supervised classification using Random Forest and XGBoost. SMOTE oversampling was applied to address class imbalance before training. Models were evaluated using Precision, Recall, F1-score, and Confusion Matrix analysis. XGBoost consistently outperformed Random Forest, achieving the best overall F1-score.

Module 4 – Explainability and Dashboard: SHAP (TreeExplainer) was integrated to provide local, feature-level interpretability for each prediction, and SHAP outputs were converted into human-readable explanations (e.g., “MRP is 1.3× higher than the estimated real market price”). An interactive dashboard was built using Streamlit and Plotly, allowing real-time product selection, 90-day price-trend visualization, MRP-vs-selling-price comparison, and display of prediction confidence with class probabilities — completing an end-to-end pipeline from database to ML models to explainability to user interface.



Fig-2: MRP vs Selling Price by predicted verdict

C. Technical Specifications

Table 1: Technology Stack

Component	Technology	Purpose
Development Env.	VS Code	IDE for scripts and notebooks
Core Development	Python	End-to-end pipeline implementation
Data Processing	Pandas, NumPy	Cleaning, transformation, feature engineering
ML Pipeline	Scikit-learn, RF, XGBoost	Model building, training, evaluation
Imbalance & Anomaly	SMOTE, Isolation Forest	Class balancing, anomaly detection
Visualization	Plotly, Streamlit	Interactive dashboard and charts
Storage	SQLite, CSV	Structured data and outputs
Model Persistence	Joblib, GitHub	Model saving and version control
Evaluation	F1, Confusion Matrix, SHAP	Performance measurement, interpretability

4. RESULTS AND DISCUSSION

A. Model Performance

The hybrid pipeline was evaluated on a held-out test set. XGBoost achieved a weighted F1-score of 0.969, outperforming Random Forest (F1 = 0.935).

Table 2: Model Comparison (Test Set)

Model	F1-Score	Notes
Random Forest	0.935	Strong baseline; some Fake↔Suspicious confusion
XGBoost	0.969	Best performer; adopted as final classifier

Table 3: XGBoost Confusion Matrix (Test Set)

Actual \ Pred.	Fake	Genuine	Suspicious
Fake	132	0	3
Genuine	0	151	2
Suspicious	5	4	153

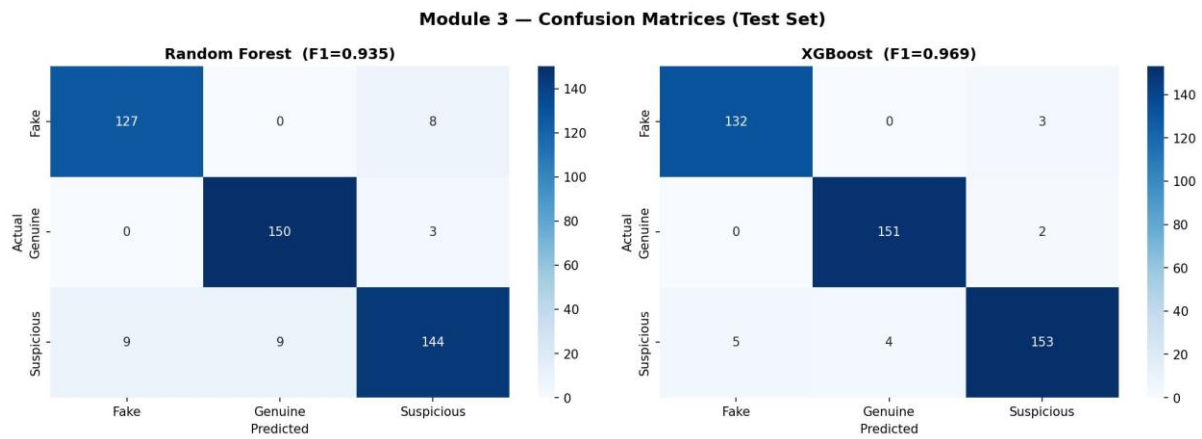


Fig-5: Confusion matrices for Random Forest and XGBoost (Test Set)

B. Feature Importance

Feature importance analysis confirmed that the composite fraud score and MRP inflation ratio were, by a wide margin, the strongest predictors of fraud, followed by discount gap magnitude, pre-sale spike detection, and seller trust score — validating the domain-driven feature engineering approach over relying on raw price fields alone.

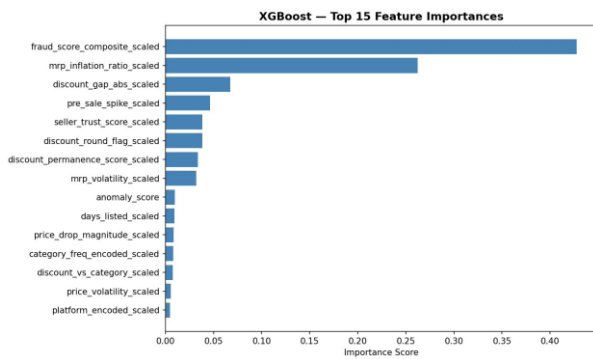


Fig-3: XGBoost – Top 15 feature importances

C. Explainability and Dashboard Output

The system does more than classify — it explains. For a sample product, SHAP output identified that the product had recorded 0% real discount over the last 90 days despite an advertised 23.6% discount, and that the listed MRP was 1.3× the estimated real market price, producing a transparent, per-product fraud score alongside class probabilities.

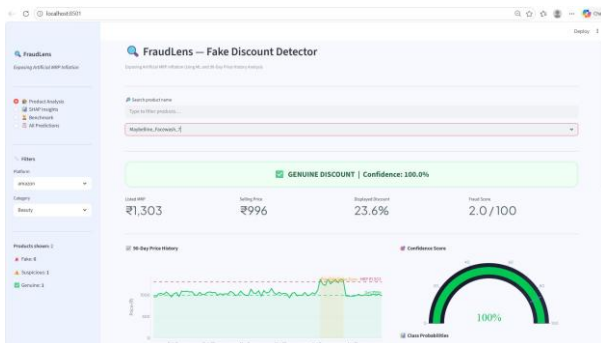


Fig-4: FraudLens interactive Streamlit dashboard output

Overall, by combining anomaly detection with supervised classification, the system not only classifies products as Fake,

Genuine, or Suspicious but also explains why — identifying artificial price hikes before discounts and making the results transparent and actionable rather than a black-box output.

5. INFERENCE

The hybrid approach (Isolation Forest + XGBoost) successfully captures both hidden anomalies through the unsupervised layer and clear fraud patterns through the supervised layer. Including the anomaly score as an input feature meaningfully improved classification performance, confirming that fraud detection in this domain is inherently multi-dimensional — no single feature is sufficient on its own. Products labeled “Suspicious” occupy a genuine gray zone and are inherently harder to classify than clear Fake or Genuine cases. SHAP-based explainability confirmed that model predictions are data-driven and interpretable rather than opaque.

6. FUTURE SCOPE

Planned extensions include a real-time browser extension to alert users about fake discounts while shopping; predictive models to estimate future price drops and suggest optimal purchase timing; secure long-term storage of historical pricing data for greater transparency; expansion to other dynamically priced domains such as travel and grocery platforms; personalized recommendations based on user behavior; a mobile application for wider accessibility; a larger-scale training dataset; and integration of more advanced AI models to further improve detection accuracy.

7. CONCLUSION

This project has designed, implemented, and validated FraudLens, a hybrid fraud detection framework that combines rule-based logic, unsupervised anomaly detection, and supervised machine learning classification to identify fake discounts in e-commerce. By anchoring every discount claim against a 90-day historical price baseline and layering Isolation Forest anomaly scores into a Random Forest and XGBoost classification pipeline, the system achieves high detection accuracy (weighted F1-score of 96.9%) while remaining interpretable through SHAP-based explainability. The accompanying interactive dashboard demonstrates that

fraud detection results can be delivered to consumers in a transparent, actionable, and easy-to-understand format, moving beyond descriptive price trackers toward a genuine diagnostic tool for pricing integrity.

ACKNOWLEDGEMENT

We express our sincere gratitude to our mentor, Mrs. Anitha M, Assistant Professor, Department of Computer Science and Engineering, Avinashilingam Institute for Home Science and Higher Education for Women, for her invaluable guidance and insightful feedback throughout this project. We also thank the Department of Computer Science and Engineering for providing the infrastructure and resources necessary to carry out this work.

REFERENCES

- [1] H. L. S. Gino, J. R. Ponciano, C. D. G. Linhares, and A. J. M. Traina, "ECommVis: Supporting E-Commerce Marketplace Advertising Outcomes Through a Visual Analytics System," *IEEE Open Journal of the Computer Society*, vol. 7, pp. 455–467, 2026.
- [2] Y. Yada, J. Feng, T. Matsumoto, N. Fukushima, F. Kido, and H. Yamana, "Dark Patterns in E-Commerce: A Dataset and Its Baseline Evaluations," in *Proc. 2022 IEEE International Conference on Big Data (IEEE BigData)*, Osaka, Japan, 2022, pp. 3015–3022, doi: 10.1109/BigData55660.2022.10020800.
- [3] J. Ramakrishnan, E. Shaabani, C. Li, and M. A. Sustik, "Anomaly Detection for an E-Commerce Pricing System," *arXiv preprint arXiv:1902.09566*, 2019.
- [4] A. Mathur, G. Acar, M. J. Friedman, E. Lucherini, J. Mayer, M. Chetty, and A. Narayanan, "Dark Patterns at Scale: Findings from a Crawl of 11K Shopping Websites," *Proc. ACM Hum.-Comput. Interact.*, vol. 3, no. CSCW, Article 81, pp. 1–32, Nov. 2019, doi: 10.1145/3359183.
- [5] Y. Xu, M. Li, L. Cui, S. Huang, F. Wei, and M. Zhou, "LayoutLM: Pre-training of Text and Layout for Document Image Understanding," in *Proc. 26th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining (KDD '20)*, 2020, pp. 1192–1200, doi: 10.1145/3394486.3403172.