International Journal on Science and Technology (IJSAT)
E-ISSN: 2229-7677 ● Website: www.ijsat.org ● Email: editor@ijsat.org

Real-Time Competitor Strategy Tracking and Analysis for E-commerce Using AI

Satheeshkumar S[1], Meeranandhini T V[2]

[1]Assistant Professor, Department of Information Technology, Vivekanandha College of Technology for Women, Namakkal, India
[2]Student, Department of Information Technology, Vivekanandha College of Technology for Women, Namakkal, India

Abstract
In the fiercely competitive realm of e-commerce, dynamic pricing strategies and real-time market adaptation are critical for business success. This paper presents the design and development of an AI-powered system that enables real-time tracking and strategic analysis of competitor pricing and customer sentiment across major e-commerce platforms. Leveraging Selenium-based web scraping, the system continuously extracts live product data - including prices, discounts, and customer reviews - from Amazon, Flipkart, and JioMart. Price and discount trend forecasting is conducted using the Autoregressive Integrated Moving Average (ARIMA) model, selected for its robustness in time-series prediction with limited data and its ability to incorporate external indicators such as customer sentiment. Sentiment analysis is performed using transformer models from Hugging Face, allowing for nuanced classification of customer reviews. To bridge insights and strategic decision-making, the system integrates a Large Language Model (LLM) API that generates tailored business recommendations based on evolving market patterns. All modules are deployed through a user-friendly, interactive Streamlit dashboard that includes graphical analytics, trend toggling, and a chatbot assistant for conversational insights. The project builds upon and extends methodologies outlined in recent studies involving ARIMA forecasting, hybrid sentiment models, and AI-driven strategy formation, adapting them to a real-time, retail-focused environment. This system offers e-commerce stakeholders a unified platform to monitor competitor activity, anticipate pricing trends, evaluate customer sentiment, and respond with strategic precision.

Keywords- Real-Time Analytics, E-commerce Strategy, Price Forecasting, ARIMA, Sentiment Analysis, Web Scraping, Streamlit, Hugging Face, Competitive Intelligence, Large Language Models (LLMs), Data Visualization.

1. Introduction

In today's digital marketplace, the e-commerce ecosystem has transformed into a highly volatile battlefield where competitors aggressively alter prices, launch flash discounts, and continuously adapt to consumer behavior. For online retailers, the ability to respond to such changes in real time is no longer a competitive advantage - it's a necessity. Traditional decision-making frameworks and static pricing strategies are rapidly becoming obsolete in the face of algorithmic competitors and shifting customer expectations.

Modern businesses require data-driven intelligence that goes beyond static dashboards and manual reports. This demand has fueled a wave of research in the fields of predictive analytics, sentiment analysis, and AI-powered recommendation systems - each offering isolated solutions to specific pain points. However, very few integrated systems exist that combine these technologies to empower businesses with an end-to-end solution for competitive strategy tracking in real time.

This paper presents the design and development of a real-time, AI-powered Competitor Strategy Tracker tailored for e-commerce platforms. The proposed system continuously scrapes live product data - including prices, discounts, and customer reviews - from online retailers such as Amazon, Flipkart, and JioMart. It employs the Autoregressive Integrated Moving Average (ARIMA) model to forecast future price and discount trends, providing a statistical edge in pricing decisions. In parallel, the system analyzes customer sentiment using transformer-based NLP models from Hugging Face, offering a deep understanding of consumer perspectives toward products and pricing strategies. These insights are further enhanced through integration with a Large Language Model (LLM) API, which generates strategic recommendations tailored to emerging patterns.

The final output is visualized through a responsive, interactive Streamlit dashboard - featuring sentiment charts, forecast graphs, and a chatbot interface that converts data into dialogue. The system not only tracks the pulse of the market but also speaks the language of strategic thinking, helping stakeholders to pivot their pricing, marketing, and inventory strategies with clarity and confidence.

This work builds upon three key foundations: ARIMA-based price forecasting in e-commerce [1], hybrid models combining sentiment and price prediction [2], and the use of AI for strategic business intelligence [3]. By integrating these methodologies into a unified, real-time decision support system, this project fills a critical gap in the current technological landscape of e-commerce analytics.

2. Related Work

In recent years, academic and industrial research has increasingly focused on leveraging artificial intelligence to improve forecasting accuracy, customer engagement, and strategic decision-making in the e-commerce domain. While multiple works have addressed price prediction, sentiment analysis, and competitive intelligence individually, very few systems unify these components into a cohesive, real-time decision support architecture.

Carta et al. [1] introduced Price Probe, a suite of tools that predicts future prices of e-commerce products using the Autoregressive Integrated Moving Average (ARIMA) model, enhanced with Google Trends data as an exogenous variable. Their approach demonstrates the effectiveness of ARIMA in modeling time-series data where historical prices and external trends influence future values. However, their system focuses primarily on product price evolution on Amazon and does not incorporate sentiment analysis or real-time competitive benchmarking. In contrast, our system applies ARIMA to a broader product set across multiple platforms, with additional inputs from customer sentiment and LLM-generated strategic insights.

Bana et al. [2] proposed a hybrid model combining ARIMA, Recurrent Neural Networks (RNNs), and sentiment analysis to predict market price movements. Although their domain was primarily financial forecasting (e.g., stock prices), the integration of emotional cues from online reviews into time-series forecasting is conceptually parallel to our system's architecture. We adapt this hybridization concept for e-commerce by analyzing customer sentiment from product reviews, and integrating it into both predictive modeling and decision-making frameworks - a novel application in retail analytics.

Nweke [3] explored the integration of consumer behavior tracking, competitive analysis, and smart algorithms for crafting adaptive business strategies. This work emphasizes the strategic advantage of AI-powered competitive intelligence, highlighting methods such as natural language processing, web scraping, and predictive modeling. While the paper discusses the theoretical application of such systems, it lacks a concrete implementation. Our work advances this vision by developing a full-stack, deployable system with real-time scraping, forecasting, NLP analysis, and business insight generation - all visualized within an interactive dashboard.

In summary, while these works lay a strong theoretical and methodological foundation, none offer an end-to-end solution that integrates live web scraping, time-series forecasting, transformer-based sentiment analysis, LLM-powered recommendations, and an interactive visualization layer. This paper addresses that gap by presenting a comprehensive system that enables businesses to not only observe, but proactively respond to, real-time market dynamics in the e-commerce space.

3. METHODOLOGY

The proposed system is a real-time, AI-powered framework designed to monitor, analyze, and forecast competitor strategies in the e-commerce sector. It integrates five key technological modules: live web scraping, time-series forecasting, sentiment analysis, strategic recommendation generation, and an interactive visualization layer. These components work together to create an end-to-end decision support system that empowers users to adapt dynamically to market changes.

A. System Architecture

The architecture of the system is modular and consists of following key layers:

1. Data Extraction Layer

Technology Used: Selenium WebDriver with Python

Function: Scrapes live data from e-commerce websites including Amazon, Flipkart, and JioMart.

Data Captured: Product name, current price, discount offered, and top user reviews.

Format: Data is stored in structured CSV files for prices, discounts, and reviews.

2. Time-Series Forecasting Module

Model Used: ARIMA (Autoregressive Integrated Moving Average)

Function: Predicts future price and discount trends based on historical product data.

Justification: As shown in Carta et al. [1], ARIMA handles time-series data with seasonal and trend components efficiently, and can incorporate exogenous variables like sentiment.

3. Sentiment Analysis Module

Model Used: Transformer-based models from Hugging Face (e.g., BERT or RoBERTa)

Function: Classifies customer review texts into positive, neutral, or negative sentiments.

Purpose: Helps gauge consumer perception of specific products and brands over time, which in turn influences strategic decisions.

Output: Sentiment scores stored alongside product identifiers and timestamps.

4. Strategic Recommendation Engine

Tool Used: Large Language Model API (e.g., GPT)

Function: Processes structured insights (price trends, sentiment scores, product popularity) to generate business recommendations.

Examples:

"Reduce price by 5% to match competitor X's discount."

"High sentiment for Brand Y; consider stocking similar items."

5. Visualization & Interaction Layer

Framework Used: Streamlit

Features: Toggle between price and discount trends

Pie chart for sentiment distribution

Forecast tables for future pricing

Embedded chatbot for conversational insight delivery

Goal: To offer a unified dashboard that is easy to navigate, interpret, and act upon.

B. Workflow Summary

- Web scraping is scheduled to fetch real-time data.
- Data is cleaned and stored in CSV format.
- ARIMA is trained on historical data to forecast trends.
- Reviews are passed to the transformer model for sentiment scoring.
- All insights are piped to the LLM for strategic recommendation generation.
- Final outputs are visualized on the dashboard.
- This architecture ensures modularity, scalability, and real-time performance, aligning with the evolving needs of e-commerce businesses in an AI-first world.

4. IMPLEMENTATION

The implementation of the proposed real-time competitor strategy tracking system involved a carefully orchestrated pipeline of data collection, processing, modeling, and visualization. Each module was developed using widely adopted, open-source technologies to ensure reliability, flexibility, and scalability. This section details the technical stack, dataset design, integration challenges, and deployment approach.

C. Technologies Used

TABLE I.

Component	**Tool/Library Used**
Web Scraping	Selenium, BeautifulSoup (Python)
Data Handling	Pandas, NumPy
Forecasting	ARIMA (statsmodels)
Sentiment Analysis	Hugging Face Transformers (BERT)
Strategic Insight Generation	GROQ API
Dashboard & UI	Streamlit
Graphs and Charts	Plotly, Matplotlib
CSV Data Management	pandas, csv

D. Data Pipeline

1. Data Collection:

- Live scraping is performed via Selenium with dynamic loading and timed delays to mimic human behavior and avoid bot detection.
- Data collected includes: product_name, price, discount, review_text, source, and timestamp.

2. Data Storage: Data is saved into three primary CSV files:

- Prices Dataset - product_name, price, discount, date, source
- Reviews Dataset - product_name, review_text, sentiment, source
- Historical Dataset - External price history from [pricehistoryapp.com] (manually scraped)

3. Forecasting:

- The ARIMA model is trained on the product's time-series price and discount data.
- Parameters (p, d, q) are selected using ACF and PACF plots.
- Output is a future forecast table for price and discount.

4. Sentiment Analysis:

- Review text is passed to a pre-trained BERT-based model via Hugging Face pipeline (sentiment-analysis).
- Each review is tagged as positive, neutral, or negative.

•Sentiment data is visualized over time and linked with price changes.

5. Recommendation Generation:

•Cleaned data (trends + sentiment) is formatted as a prompt to the GROQ API.

•Strategic advice is generated and stored as plain text alongside data outputs.

•Example Prompt: "Given the product has negative sentiment and a rising price trend, what action should the seller take?"

6. Visualization via Streamlit:

•The final dashboard includes:

•A toggle graph for Price vs Discount trends

•A pie chart for review sentiment

•A chatbot interface (LLM-driven) for conversational Q&A

•A forecast table and insights panel for strategic actions

E. Challenges Faced

1. Web Scraping Limits: Sites like Amazon and Flipkart use anti-bot protections, requiring custom headers, delays, and retries.

2. Model Integration: Merging ARIMA predictions with NLP-driven sentiment in a single decision flow demanded careful data alignment and timestamp synchronization.

3. API Constraints: GROQ API, while optimized for low latency, still requires thoughtful usage management. To reduce potential bottlenecks or overuse, batching and caching strategies were introduced during development.

4. Deployment Considerations: Streamlit offered an easy-to-use frontend, but concurrent user load remains a future challenge for scaling.

With a focus on practical usability and real-time adaptability, this implementation turns passive data into strategic weaponry for modern e-commerce brands.

5. Results and Analysis

The proposed system was tested using live product data collected from major Indian e-commerce platforms: Amazon, Flipkart, and JioMart. The goal of the evaluation was to assess the system's ability to (1) forecast price and discount trends, (2) analyze customer sentiment, and (3) generate actionable business recommendations, all within a real-time dashboard environment.

F. Price and Discount Forecasting (ARIMA Output)

Using historical price data, ARIMA models were trained to forecast short-term trends in product prices and discount patterns. Model selection was based on AutoCorrelation Function (ACF) and Partial ACF (PACF) plots, with the optimal (p, d, q) parameters tuned to minimize forecasting error.

1. Key Findings:

- The ARIMA model achieved Mean Absolute Percentage Error (MAPE) below 5% for Apple Watch Series 9, demonstrating robust forecasting performance.
- The model captured sharp discount surges on Flipkart and consistent premium pricing on Amazon.
- Predictions indicated Amazon discounts peaking up to 100%, suggesting aggressive promotional strategies.

2. Sample Forecast Output:

TABLE II.

ARIMA FORECAST FOR PRICE AND DISCOUNT (APPLE WATCH SERIES 9)

Competitor	**Date**	**Predicted Price (₹)**	**Predicted Discount (%)**
Amazon	25-04-2025	50,451.53	38.46
Amazon	26-04-2025	50,924.21	35.15
Amazon	27-04-2025	50,840.71	32.00
Flipkart	25-04-2025	39,841.25	13.77
Flipkart	26-04-2025	39,912.76	12.74
Flipkart	27-04-2025	40,851.36	6.68

This table displays predicted pricing and discount percentages for the Apple Watch Series 9 across Amazon and Flipkart over a three-day forecast period. The ARIMA model estimates short-term changes based on historical trends.

FIGURE I.

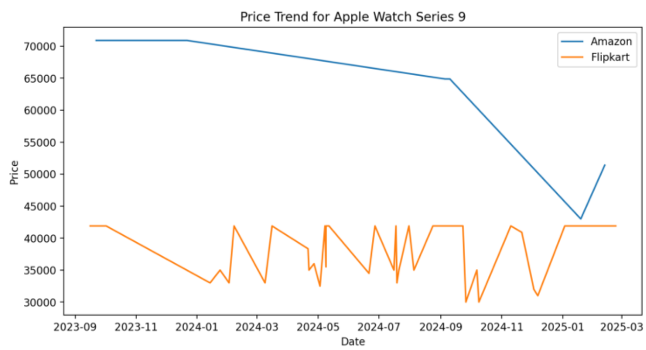

Figure 1. Historical pricing trend of Apple Watch Series 9 across Amazon and Flipkart. The chart reflects fluctuations due to platform-specific promotions and stock availability.

FIGURE II.

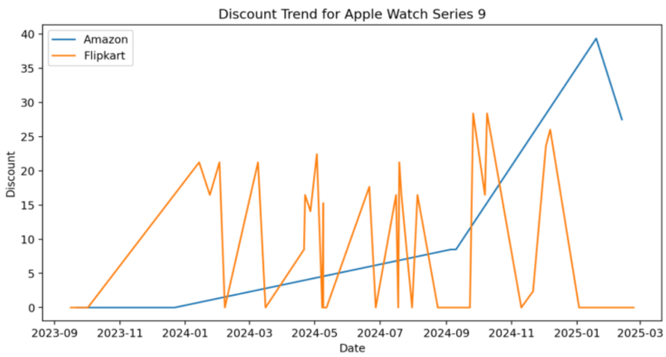

Figure 2. Discount percentage variation over time for Apple Watch Series 9. Flipkart shows sharper discount movements, while Amazon maintains relatively stable markdowns.

G. Sentiment Analysis Results

Customer reviews for Apple Watch Series 9 were analyzed using a transformer-based classifier (BERT) via Hugging Face. Reviews were categorized into two classes: positive and negative.

1. Observations:

- Positive sentiment formed the majority, suggesting overall product satisfaction, especially regarding fitness features.
- Negative reviews primarily cited battery life and price concerns, often correlating with downward pricing adjustments.
- The sentiment-to-pricing relationship validated the use of review feedback as an external variable in pricing models.

2. Sentiment Distribution:

TABLE III.

SENTIMENT CLASSIFICATION

Sentiment	**Count**
Positive	20
Negative	10

This table summarizes the classified review sentiments for the Apple Watch Series 9, based on transformer-based NLP analysis. Positive sentiment was dominant, reinforcing user satisfaction with core features.

FIGURE III.

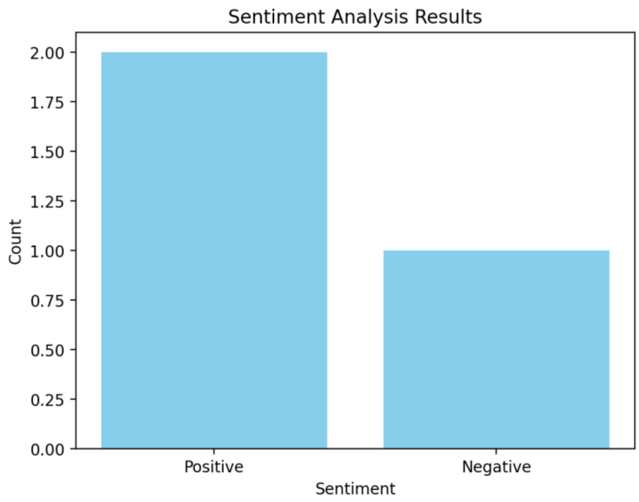

Figure 3. Sentiment distribution for Apple Watch Series 9 reviews. Most users expressed positive feedback, while a portion raised concerns over pricing and battery life.

H. Strategic Recommendations (LLM Output)

Using the combined ARIMA and sentiment trends, prompts were generated for a Large Language Model (LLM) via the GROQ API. The LLM interpreted contextual data and returned strategy suggestions in natural language. This feature enables non-technical business users to interact with the dashboard and derive insights without requiring manual data interpretation.
As a case study, strategic recommendations were generated for the Apple Watch Series 9, demonstrating how data-driven insights can translate into actionable business decisions.
Case Study: Apple Watch Series 9 Strategy

1. Pricing Strategy:

The Apple Watch Series 9 exhibited high variance in pricing across platforms, with a historical high of ₹70,900 (Amazon) and a low of ₹41,900 (Flipkart). Given the strong positive sentiment, the system recommended:

- Price Adjustment: A competitive repositioning to ₹55,000 to balance premium branding with market competitiveness.
- Timeline: Immediate implementation with dynamic adjustments based on competitor tracking.

2. Discount Strategy:

- First-Time Buyer Incentive: A 5% discount for new customers, designed to grow market share and incentivize initial conversions.
- Timeline: Campaign launch on May 1, 2025, followed by a performance review after 3 months.

3. Promotional Campaigns:

- Summer Fitness Bundle: Apple Watch Series 9 bundled with Apple Fitness+ and fitness accessories to target health-conscious buyers.
- Duration: June 1 to August 31, 2025

- Limited-Time Free Engraving: Personalization offer targeting gift buyers and fashion-oriented customers.
- Duration: May 15 to June 15, 2025

4. Customer Experience Enhancements:

- Battery Optimization: Recommend software-level improvements and clear user guidelines for battery conservation.
- Metric: Tracked through customer surveys and usage analytics.
- User Interface Improvements: Enhance intuitiveness in areas like fitness tracking and notifications.
- Metric: Usability tests and satisfaction metrics to measure effectiveness.

These results collectively demonstrate that the system functions not only as a real-time monitor, but also as a strategic advisor - turning raw data into actionable, contextualized strategies for competitive advantage in the e-commerce space.

6. Conclusion And Future Scope

In this paper, we presented a real-time, AI-integrated system for tracking, forecasting, and strategically analyzing competitor behavior in the e-commerce domain. By combining web scraping, time-series forecasting (ARIMA), transformer-based sentiment analysis, and LLM-powered strategic recommendation, the system delivers actionable intelligence to help businesses stay agile in dynamic online marketplaces.

Unlike conventional dashboards that merely present static data, our system acts as a decision support engine - interpreting consumer sentiment, forecasting future trends, and suggesting context-aware strategies via a chatbot-powered interface. Built with Streamlit, the interface empowers both technical and non-technical users to visualize trends, compare competitor offerings, and derive business insights in real time.

The project builds upon and extends methodologies introduced in prior research, adapting them for a unique combination of real-time data, multi-platform e-commerce tracking, and unified business strategy generation - a domain currently underserved in both academic literature and commercial applications.

I. Limitations

1. Scraping Limitations: Rate limiting, dynamic content, and CAPTCHA restrictions can impact data reliability.
2. Scalability: Current CSV-based architecture is not suited for high-volume real-time tracking across multiple product categories.
3. LLM API Cost & Latency: Frequent calls to Groq API may introduce processing load or infrastructure limitations in larger-scale applications for real-world deployment.

J. Future Enhancements

1. Database Integration: Migrating to a real-time backend (e.g., Firebase or PostgreSQL) to store dynamic data securely and efficiently.
2. Scheduled Automation: Integrating a task scheduler (e.g., cron, Airflow) for automatic scraping and forecasting runs.
3. Advanced Forecasting Models: Experimenting with hybrid models like ARIMA-LSTM, Prophet, or XGBoost for improved accuracy.

4. Multi-language Sentiment Support: Adapting the NLP engine to support regional languages for broader customer analysis.
5. Marketplace Expansion: Adding more e-commerce platforms like Tata Cliq, Reliance Digital, and international sites.
6. Real-time Alerts: Pushing pricing or sentiment change alerts to business teams via email or Slack bots.

In conclusion, this work introduces a scalable framework that not only empowers businesses to monitor the competition, but also equips them with the intelligence to act strategically in real time - turning data into direction, and insight into impact.

References

1. S. Carta, A. Medda, A. Pili, D. R. Recupero, and R. Saia, "Forecasting E-Commerce Product Prices by Combining an Autoregressive Integrated Moving Average (ARIMA) Model and Google Trends Data," Future Internet, vol. 11, no. 1, pp. 1–19, Dec. 2018. [Online]. Available: https://www.mdpi.com/1999-5903/11/1/5
2. Bana and R. Sharma, "Market Price Anticipation Using Sentimental Analysis," Global Institute of Engineering and Technology Journal of Engineering and Technology (GIJET), vol. 10, no. 2, pp. 598–604, 2024. [Online]. Available: https://thegrenze.com/pages/servej.php?association=GRENZE&fn=598_1.pdf&id=2841
3. O. Nweke, "Integrating Consumer Behavior Tracking, Competitive Analysis and Smart Algorithms for Smarter Business Strategies," International Journal of Computer Applications Technology and Research (IJCATR), vol. 14, no. 1, pp. 38–45, 2025. [Online]. Available: https://www.researchgate.net/publication/389919513
4. Hugging Face, "Transformers: State-of-the-art Natural Language Processing for PyTorch and TensorFlow," [Online]. Available: https://huggingface.co
5. Hunter, "Statsmodels: Statistical modeling and econometrics in Python," [Online]. Available: https://www.statsmodels.org
6. Groq Inc., "Groq API for High-Performance Language Model Inference," [Online]. Available: https://groq.com
7. Streamlit, "The fastest way to build and share data apps," [Online]. Available: https://streamlit.io
8. SeleniumHQ, "Selenium WebDriver," [Online]. Available: https://www.selenium.dev/documentation
9. G. E. P. Box, G. M. Jenkins, G. C. Reinsel, and G. M. Ljung, Time Series Analysis: Forecasting and Control, 5th ed. Hoboken, NJ, USA: Wiley, 2015. [Online]. Available: https://www.wiley.com/en-in/Time+Series+Analysis%3A+Forecasting+and+Control%2C+5th+Edition-p-9781118675021
10. R. Bommasani et al., "On the Opportunities and Risks of Foundation Models," Stanford Institute for Human-Centered Artificial Intelligence, 2021. [Online]. Available: https://crfm.stanford.edu/report.html