

# DemandCast – AI-Based Demand Forecasting for E-commerce Products

---

## Project Domain & Purpose

### Domain:

- Retail & E-commerce
- Machine Learning (ML) & Time Series Forecasting
- Natural Language Processing (NLP) & Social Media Analytics
- Supply Chain Optimization

### Purpose:

As e-commerce continues to expand, accurate demand forecasting is essential for optimal stock management, reduced waste, and responsive marketing. Most existing solutions focus solely on past sales, ignoring external factors (social sentiment, promotional events, search trends) that dynamically affect demand. This project fills that gap by developing a robust forecasting system that brings together traditional time series models, advanced ML, and external data integration for business decision support.

---

## Problem Statement

Traditional forecasting models (like ARIMA or classical regression) often fail to capture abrupt demand shifts triggered by marketing campaigns, social media buzz, or seasonal festivals. These blind spots lead to overstocking (increased holding costs) or stockouts (lost sales/opportunity). A smarter, hybrid system is needed to recognize and react to these real-world influences on demand.

---

## Project Objectives

- **Predict product-level (SKU/category) demand** on a daily/weekly horizon.
- **Capture historical sales patterns:** Trend, cyclic seasonality, calendar events.
- **Integrate external variables:**
  - *Google Trends:* Real-time search intent.
  - *Twitter sentiment:* Public opinion/consumer buzz (analyzed via VADER/BERT).
  - *Promotional calendar:* Sale periods, discounts, festival spikes.
- **Model and compare methodologies:**
  - Classical time series (ARIMA, Prophet), ML regression (XGBoost), sequential deep learning (LSTM).
- **Business-facing dashboard:** Input/upload sales data, configure external features, visualize forecasted demand, uncertainties, and explanatory trends.

# Proposed Solution & Methodology

## 1. Data Acquisition & Preparation

- **Historical sales:** Collect open datasets (e.g., Kaggle e-commerce, UCI retail).
- **External signals:**
  - *Google Trends & Twitter:* Use API wrappers to fetch search frequency and tweets mentioning product/category.
  - *Sentiment Analysis:* Apply VADER for rule-based, BERT for deep contextual analysis.
  - *Event tagging:* Curate promotional/festival dates manually (CSV/Excel template).
- **Data cleaning:** Handle missing values, outlier detection, calendar alignment.

## 2. Feature Engineering

- **Time Features:** Day-of-week, month, lags (previous sales), rolling averages.
- **Event Features:** Binary flags or impact weights for promo/festival periods.
- **Sentiment Scores:** Convert daily/weekly text signals to numerical scores usable in forecasting models.
- **Search Trends:** Normalize Google Trends data to match sales volume scales.

## 3. Model Development & Benchmarking

- **Statistical:**
  - *ARIMA:* Models autocorrelation/trend, best for univariate time series.
  - *Prophet:* Handles seasonality, holidays, external regressors seamlessly.
- **Machine Learning:**
  - *XGBoost/LightGBM:* Ensemble regression using both time and external features (lags, sentiment, search, event flags).
- **Deep Learning:**
  - *LSTM:* Sequence-to-sequence learning incorporating both sales and external features (requires larger datasets).
- **Model Evaluation:** Calculate RMSE, MAE, MAPE metrics. Perform cross-validation, and report best-performer for each product/category.

## 4. Dashboard Design & Visualization

- *Streamlit/Flask-based app:*
  - Upload sales data and configure external regressors.
  - View predicted demand, confidence intervals, and historical vs. forecasted lines.
  - Explore breakdowns by campaign impact, festival effect, sentiment trend overlays.

# Tools & Technologies

Type	Tools/Libraries
Data Analysis	Python (Pandas, NumPy)
Time Series Modeling	ARIMA, Prophet
ML/DL Modeling	XGBoost, LightGBM, LSTM (TensorFlow)
NLP/Sentiment Analysis	VADER, BERT
Trend Data	Google Trends API
Social Media	Twitter API, Tweepy, SNScrape
Visualization	Matplotlib, Plotly, Streamlit

---

## Deliverables

- End-to-end demand forecasting system (source code, API scripts, documentation).
- Interactive dashboard – demonstrate uploads, forecasts, explanatory graphs.
- Thorough report (PDF) outlining methodology, technical choices, evaluation results, and business implications.
- (Optional) GitHub repository and YouTube/video demo for public showcase.

---

## Significance & Impact

- Provides actionable forecasts for marketing, pricing, and procurement teams.
- Reduces operational inefficiencies by anticipating demand spikes/slumps due to external events.
- Methods generalize to other retail, supply chain, and even healthcare domains.
- Demonstrates integration of classical analytics, machine learning, NLP, and business intelligence.

---

## Conclusion

The **DemandCast** project delivers a comprehensive solution for modern, feature-rich demand forecasting, bridging analytics, machine learning, and real-world external signals. It stands out for its innovation, practicality, and direct value for e-commerce analytics and business strategy roles.

---