

Application of AI in E-Commerce for Startups with  
Limited Data: A Lean Startup Approach to Demand  
Forecasting and Inventory Optimization

MSc Research Project  
MSc Artificial Intelligence for Business

Ana Paula Elizondo de la Garza  
Student ID: 23140879

School of Computing  
National College of Ireland

Supervisor: Victor del Rosal

**National College of Ireland**  
**MSc Project Submission Sheet**



**School of Computing**

**Student Name:** Ana Paula Elizondo de la Garza  
**Student ID:** 23140879  
**Programme:** MSc A.I. for Business **Year:** 2025  
**Module:** Practicum  
**Supervisor:** Victor del Rosal  
**Submission Due Date:** August 11, 2025  
**Project Title:** Application of AI in E-Commerce for Startups with Limited Data: A Learn Startup Approach to Demand Forecasting and Inventory Optimization  
**Word Count:** 5,890 words  
**Page Count:** 25 pages

I hereby certify that the information contained in this (my submission) is information pertaining to research I conducted for this project. All information other than my own contribution will be fully referenced and listed in the relevant bibliography section at the rear of the project.  
ALL internet material must be referenced in the bibliography section. Students are required to use the Referencing Standard specified in the report template. To use other author's written or electronic work is illegal (plagiarism) and may result in disciplinary action.

**Signature:** Ana Paula Elizondo de la Garza

**Date:** August 10, 2025

**PLEASE READ THE FOLLOWING INSTRUCTIONS AND CHECKLIST**

Attach a completed copy of this sheet to each project (including multiple copies)	<input type="checkbox"/>
<b>Attach a Moodle submission receipt of the online project submission,</b> to each project (including multiple copies).	<input type="checkbox"/>
<b>You must ensure that you retain a HARD COPY of the project,</b> both for your own reference and in case a project is lost or mislaid. It is not sufficient to keep a copy on computer.	<input type="checkbox"/>

Assignments that are submitted to the Programme Coordinator Office must be placed into the assignment box located outside the office.

<b>Office Use Only</b>	
Signature:	
Date:	
Penalty Applied (if applicable):	

# Application of AI in E-Commerce for Startups with Limited Data: A Lean Startup Approach to Demand Forecasting and Inventory Optimization

Ana Paula Elizondo de la Garza  
23140879

## Abstract

Startups in the e-commerce sector face critical operational challenges due to data-scarcity, inventory volatility and limited resources, which significantly affects demand forecasting accuracy and efficient inventory planning. This research explores how AI-based solutions can benefit startups in these specific contexts by designing and evaluating a lightweight, startup-ready forecasting system. This study is tested with real transactional data from a Mexican perfumes e-commerce startup, LeBorêt, and compared four predictive models: SARIMAX with bootstrapping, baseline LSTM, enhanced LSTM with data augmentation and transfer learning, and XGBoost with fine-tuned features. Evaluation results shows that while LSTM-models offer strong adaptability, SARIMAX combined with bootstrapping provided the most operationally significant balance between accuracy levels and flexibility, achieving 93.3% coverage in a predictive interval. The selected forecasting method was integrated into a minimum viable product (MVP), “COSMOS”, following the Lean Startup methodology focusing on validated learning by co-designing this interface with the startup’s founder. The system offers dynamic forecasts, smart inventory alerts and various planning tools. This project contributes to AI democratization, and offers a validated, replicable methodology for AI adoption in early-stage startups with resource-constrained situations.

## 1 Introduction

Forecasting product demand is a core function for businesses of all sizes operating in the digital economy. Nevertheless, for early-stage e-commerce startups, this represents a significant and uniquely difficult challenge. These companies often operate with limited data, small teams, constrained resources and infrastructure, yet they must respond quickly and accurately to dynamic market conditions, seasonal peaks and changing customer preferences to remain competitive. Inaccurate sales predictions can quickly lead to stockouts or overstocking situations, which damage both, customer experience and operational efficiency. Traditional forecasting methods often fall short in effectively capturing SKU-level demand fluctuations unless they are combined with segmentation strategies (Zhao, Y., 2024). Despite this, they remained the most widely used across industries until recent years, when AI-driven solutions began to gain exponential popularity for their outperforming results. However, startups rarely have the sufficient infrastructure and budget to integrate real-time forecasting into their daily operations (Lv, J., 2024).

Implementing advanced forecasting techniques such as LSTM and XGBoost models have shown significant improvements in handling these kinds of complexities, as shown by Sharma et al. (2024) and Manos et al. (2024), but the implementation of these models in a scarce-data,

high-uncertainty environments remain a gap in both research and practice. This is particularly critical for startups because being able to deliver high-quality products on time to early customers can be a decisive factor in building trust with them, gaining traction and shaping the company's future. These companies not only need accurate forecasting tools, but also lightweight, adaptable, user-friendly and cost-efficient. In this project, we address the gap by designing and testing a demand forecasting system that is specifically tailored to the real operational conditions of startups, focusing on practical AI adoption without requiring large-scale inversions, long historical data or large-scale infrastructure.

The solution is built and tested in a real Mexican startup that launched its first men's perfume line of three products – Tundra, Manhattan and Sahara – in different presentations in July 2024. [Leborêt](#) shows strong growth potential and increasing market interest, reaching 1 million Mexican pesos in sales in just 8 months since its launching campaign (A. P. Elizondo, personal communication, April 21, 2025). Nevertheless, the company is currently facing critical operation challenges. Limited data and resources have led to prolonged stockout periods, slowing down its success, and decreasing potential customer rates at a crucial moment that could determine its future. The project has three main objectives:

1. Develop an AI-driven model adapted to an e-commerce startup's environment.
2. Validate its forecasting accuracy and inventory management improvements.
3. Demonstrate its relevance and flexibility to implement in similar startups for scalability.

This study is guided by the following **research question**: How can an e-commerce startup with limited data implement AI to improve demand forecasting and prevent stockouts without significantly increasing costs? The underlying hypothesis is that advanced models such as LSTM and XGBoost will outperform traditional time series methods, SARIMAX, in terms of predictive accuracy when applied to fragmented, limited data conditions. The expected contribution of this project includes an accurate forecasting model designed for early-stage e-commerce startups, combined with a user-friendly MVP interface to standardize, automate and consolidate key tasks and next steps. This allows startups to focus on strategic planning while adopting AI-driven solutions from the beginning, helping democratize AI by reducing forecasting errors and stockouts.

## 2 Related Work

Precision in demand prediction and stockout prevention are core challenges for startups in e-commerce due to data scarcity, irregular time series and constrained operational capacities. This literature review analyses existing strategies for time series forecasting with limited data (1), identifies how transfer learning and data augmentation improve model performance in such contexts (2), explore real-world implementation of AI in startups and SMEs (3), and examines how the Lean Startup methodology provides a suitable framework for startups with uncertain environments (4).

### 2.1 Forecasting Models with Limited Data: SARIMAX, XGBoost and LSTM

SARIMAX has shown high accuracy results when the data reflect seasonal and exogenous influences. In a study performed by Jia and Yan (2023) they reached a 95.86% accuracy for daily forecasts using SARIMAX on multiregional product data with price and sales channel as exogenous variables. In this case, the model required only classical preprocessing and outperformed other models under fine time granularity. Nevertheless, it is important to highlight that SARIMAX has shown a limited capacity to model nonlinear relationships, which makes it less suitable in high volatile contexts (Özen et al., 2019) (Sharma et al., 2024). Sheng and Jia (2020) integrated SARIMAX with LSTM, this hybrid model captured both linear and nonlinear temporal patterns. Their combined model permitted users to select up to 95% confidence intervals, allowing flexible fulfillment strategies depending on risk tolerance and budget constraints.

LSTM, a deep learning model, consistently outperforms models such as ARIMA and SARIMAX in capturing complex time dependencies. Siami-Namini et al. (2018) highlighted that LSTM reduced RMSE drastically across several macroeconomic time series. Likewise, Wang and Mi (2020) demonstrated that LSTM has superior accuracy in forecasting car sharing demand under irregular seasonal patterns and trends. Similarly, in an e-commerce context, Singh et al. (2024) applied LSTM to a dataset of 4,500 customer and inventory records, achieving high performance at learning from heterogeneous inputs, although it required significantly higher training complexity. Additionally, hybrid approaches such as LSTM + Random Forest (Jessica et al., 2024) improved generalization notably, balancing accuracy and model robustness in product stock expenditure in retail forecasting. Despite its effectiveness, it is important to highlight that LSTM requires extensive hyperparameter tuning and computational resources, key considerations that may not be feasible for early-stage startups without the proper infrastructure and resources (Agarwal et al., 2021) (Mao et al., 2022). Moreover, this type of model tends to overfit on small datasets, which remains a significant concern unless properly mitigated.

On the other hand, XGBoost models represent a strong alternative for limited and noisy data contexts. Sharma et al. (2024) implemented this model for stock volatility forecasting, reporting an  $R^2$  of 0.98 and RMSE of 0.015, outperforming both LSTM and ARIMA on the same dataset. Its strength lies in its robustness against overfitting, structured data handling, engineering lagged features and training efficiently even with small datasets. Likewise, Ozen

et al. (2019) found XGBoost more reliable than SARIMA and LSTM when it was compared in the forecasting performance on multiple public time series datasets. Despite XGBoost being less expressive in long temporal dependencies, its speed, robustness and interpretability makes it particularly attractive for e-commerce applications with constrained resources.

In short, traditional statistical models such as SARIMAX offer simplicity and interpretability, while more advanced and recent approaches such as XGBoost and LSTM offer powerful pattern recognition and flexibility in scarce data environments. However, no model is universally superior, it depends on the application. The literature clearly indicates a growing trend toward hybrid models (SARIMAX-LSTM (Sheng et al., 2020), LSTM-RF (Jessica et al., 2024)) to leverage on the complementary strengths of each approach. Still, few studies explicitly evaluate these models under startup-specific conditions, such as limited and discontinuous data, real-world noise, stock volatility and shifting customer behaviour, highlighting a clear research gap. This study aims to address that gap by investigating model performance in this specific challenging and underexplored scenarios.

## **2.2 Transfer Learning and Data Augmentation**

In recent studies, transfer learning and data augmentation have proven effective in data scarce-environments in improving overall forecasting performance. These approaches are particularly relevant for startups and SMEs, which often struggle with limited data and resources. Chang and Huang (2024) applied transfer learning with ResNet50 for retail production recognition on ImageNet. In their research, they achieved up to 99.6% accuracy in their fine-tuned model while reducing training time by 76%, which reinforces this method's efficiency in dynamic retail environments with frequent product updates.

Moreover, Hirt et al. (2020) applied transfer learning with additive regression (Prophet) in sales forecasting in six different restaurant branches. In their results, they obtained a reduced MAPE by 58.7% versus local baselines, showing strong performance. On the data augmentation side, Li et al. (2024) improved small-scale time series using window slicing, scaling and noise injection, and in their results they improved RMSE by 10% in a TCN-Transformer architecture. This past study, demonstrate the value of simple augmentations for improving model generalizations. Another research by Yang et al. (2021) introduced Dynamic Time Wrapping combined with oversampling (DTW-SMOTE) to generate multiple synthetic time series. In total, they trained 17 datasets, which showed consistent accuracy gains. Nevertheless, they highlighted the significantly increased costs in computational power.

In summary, transfer learning enables rapid adaptation in cold-start adaptations while data augmentation improves overall model performance with limited but existing data.

## **2.3 Real-World AI Applications in Startups and SMEs**

AI adoption in Small and Medium Enterprises (SMEs) is advancing exponentially. In recent years, several studies showcase real-world impact: Lawton et al. (2023) introduced GM AI Foundry which enabled 30 SMEs to improve their digital maturity through innovative practices involving responsible AI, explainability and trust-building. Another study performed by

Carayannis et al. (2024) highlighted how GenAI upgraded resilience and cost-efficiency in multiple SMEs. Rajendran et al. (2024) proposed an AI strategy model for startups to define key factors such as market fit, pricing and scaling. Additionally, through several simulations to evaluate a comprehensive AI integration framework developed by Abuzaid and Alsbou showed >90% efficiency and 80% training impact for startups. These studies and research confirm that with structured frameworks, AI can significantly improve startup performance and competitiveness, even under limited resources constraints.

## **2.4 Lean Startup Methodology in AI Projects**

Eric Ries (2011) introduced in his book the Lean Startup approach which emphasizes a particular framework designed for startups and their specific environment. This framework includes rapid hypothesis testing, MVP validation, and iterative learning to reduce market risks and optimize resource use, which are highly applicable to AI-driven innovation under uncertainty scenarios. Verentennikova and Vaskiv (2018) demonstrated how the Lean Startup approach integrates with Agile and Customer Development to manage effectively innovative products. In their research they highlight the benefits on developing MVPs to test key assumptions, minimize waste and adapt features using exclusively validated learning from real customer feedback. This model reinforces iterative market testing following resource availability to enable startups to pivot quickly before scaling with higher uncertainties. Additionally, they show how low-cost experiments validate product-market fit before making considerable investments without real-world usability proof.

The reviewed literature highlights the benefits of AI integration in forecasting, data adaptation and implementation strategies for startups. Models like XGBoost and LSTM have been consistently outperforming classical statistical methods. Additionally, real-world cases studies confirm that startups can integrate AI successfully if it is supported with the right frameworks, tools and methodologies to execute it. However, several critical gaps remain in the industry: most forecasting studies use clean, organized datasets, not fragmented and scarce SKU-level datasets, typical in early-stage startups (1). Secondly, just few works evaluate forecasting models in real startup environments with shifting demand, stock variability periods and/or limited infrastructure (2). Thirdly, transfer learning applications in time-series scenarios remain unexplored, especially in contexts such as micro-retail and startups (3). Finally, while the Lean Startup methodology offers a clear guidance on project implementation, practical guidance on MVPs for AI systems is limited, particularly for inventory and sales prediction scenarios (4). This project addresses these gaps by building and testing a forecasting model tailored for resource-constrained ecommerce startups. In short, this project aims to contribute new findings regarding the intersection of AI, startup strategy and low-data innovation.

### 3 Research Methodology

E-commerce startups such as LeBorêt often struggle with limited historical data, volatile and inconsistent inventory levels, and labour-intensive manual processes that sabotage the reliability of demand forecasts and inventory control. To address these challenges in the industry, this project proposes a modular, AI-driven architecture that integrates advanced forecasting models with a user front-end interface.

The core of this project's methodology is a quantitative, comparative forecasting study designed to target the most accurate and robust model family for data-scarce retail environments. Our initial comparative proposal established in Part 1 (comparison of three configurations of LSTM: baseline, s. LSTM + data augmentation vs. LSTM + data augmentation + transfer learning) was modified for this final version, these are the updated selected models for comparison:

- **SARIMAX + Bootstrap**, to capture trends/seasonality and quantifying uncertainty.
- **XGBoost with lagged features** to capture complex, non-linear dependencies in past sales data.
- **LSTM baseline**, serving as a reference point for sequential learning.
- **LSTM + data augmentation + transfer learning**, which enhances model generalization under scarce-data conditions by pretraining on public retail series and synthetically expanding the training set.

By combining classical statistical techniques (SARIMAX) with machine learning (XGBoost) and deep-learning variants (LSTM), we can balance interpretability and adaptivity, which are key factors for a real-world e-commerce prototype.

Complementing the quantitative analysis, the front-end design follows the lean startup principles, which are incorporating continuous, semi-structured feedback from LeBorêt's founder. These feedback sessions focused on establishing feature prioritization, refinement and expansion. This validated learning resulted in a front-end MVP named "COSMOS" that includes the following:

- Forecast visualizations with evaluation metrics
- Smart-inventory recommendations and customizable alarms/alerts
- Integrated planning calendar and task-management board
- Partner management, document repository, and user settings

This iterative user-driven process to structure the front-end of this project ensures that the final system not only excels in predictive performance but also align to the real operational workflows and decision-making needs of a growing e-commerce business.

### 3.1 Data Requirements and Sources

The following list contains all the required data and sources to perform the project:

- **Primary Transactions:** Daily sales from LeBorêt’s Shopify API (2024 – 2025), including stock-out period’s flags and total sessions per day.
- **External Series:** Public perfume retail time-series from Kaggle ([“Perfume E-Commerce Dataset 2024”](#)) for LSTM pretraining and augmentations.
- **Exogenous Variables:** Holiday calendars, marketing campaign logs and day-of-week indicators.
- **User-Experience Data:** Feedback sessions (verbal feature ranking) with the Founder to build front-end feature set with validated learning.

The selected datasets were cleaned (missing-value imputation and normalization), anonymized due to a confidentiality agreement and versioned.

### 3.2 Procedures and Experimental Setup

#### 1. Data Preprocessing

- Automated Shopify data extraction of total sales per SKU, total sessions and conversion rate per day, and inventory per day and month data.
- Data cleaning by filling missing dates, stock-out tagging (and validating with Founder) and feature engineering.
- Generation of lagged variables (t-1, t-7), holiday and campaign dummies.

#### 2. Time-Aware Splitting

- Rolling-window train/validation/test to reflect genuine forecasting scenarios with the available data.

#### 3. Model Training Pipelines

- **SARIMAX + Bootstrap:** Select (p,d,q)(P,D,Q)s via AIC/BIC, then apply block-bootstrap for 95% confidence intervals.
- **XGBoost with Lags:** Grid/Bayesian search over tree and learning-rate hyperparameters.
- **LSTM Baseline:** Two-layer LSTM.
- **LSTM + Transfer Learning & Augmentation:** Pretrain on Kaggle’s selected dataset; fine-tune with LeBorêt data augmented by noise injection, time-warping, and synthetic bootstrapping.

### 3.3 Techniques and Methods Applied

**Table 1.** Chosen Techniques and Methods for the Project

Technique/Method	Description
<b>Statistical Forecasting</b>	SARIMAX for trend/seasonality, bootstrap for uncertainty quantification.
<b>Machine Learning</b>	XGBoost to capture complex, non-linear relationships among lagged features.
<b>Deep Learning</b>	LSTM variants to capture long-term dependencies; transfer learning to leverage external data.
<b>Qualitative UX Methods</b>	Conducted two usability sessions with Leborêt’s Founder. Gathered feedback on feature relevance and UI clarity through ranking exercises.

By testing these models in a data-scarce e-commerce setting, we not only determine the best forecasting strategy for Leborêt but also establish a replicable baseline for other startups facing similar challenges in the industry.

### 3.4 Data Analysis and Validation

To evaluate the effectiveness of the proposed models under scarce data environments, two standard metrics were selected: Mean Absolute Error (MAE) and Root Mean Squared Error (RMSE). Additionally, for the SARIMAX model a 95% prediction interval coverage was calculated as an extra indicator for decision-making in uncertainty situations. To simulate real-world forecasting scenarios in this project, a time-aware rolling-window approach was implemented.

### 3.5 Limitations and Ethical Considerations

#### Limitations

- **Irregular Sales Patterns & Data Gaps:** The dataset contains limited historical data of sales and stock-out periods, which can introduce bias and reduce the robustness of model fitting.
- **Transfer-Learning Constraints:** The success of transfer learning relies on the resemblance between public Kaggle time-series and Leborêt’s actual sales data. Public datasets often represent large, established retailers rather than early-stage startups operating in more volatile markets, which can potentially limit the relevance of pretrained features.

#### Ethical Considerations

- **Data Privacy:** No customer personal information is ever stored. Shopify API credentials are secured in environment variables; published data will be anonymized and provided as a consolidated CSV without sensitive fields.
- **Transparency:** All source datasets, preprocessing steps, and model hyperparameters are fully documented.
- **Human-in-the-Loop:** Final inventory recommendations, alerts, and fulfilment decisions fully rely in the founder, ensuring human oversight of all AI-driven suggestions.

- **Reproducibility:** The complete code, data-processing scripts, and model pipelines will be made publicly available ([GitHub](#)) and in the formal report.
- **Responsible Use:** The system is designed as a decision-support tool, not a decision-making authority.

## 4 Design Specification

All the data loading, preprocessing, modelling and evaluation in this project are carried out in Jupyter notebooks, Visual Studio Code and Google Collab (for TensorFlow application), each dedicated to one modelling approach. The raw data input for the entire process is the `combined_final.csv` file, which was anonymized and version controlled, containing daily records of `total_sales`, `total_sessions`, `inventory` and `stock-out flag` (ignore). In production this information would originate from the Shopify API but for confidentiality agreements, a standalone CSV file is processed.

Each notebook runs under Python 3.19 and leverages standard data-science libraries: **pandas** for data manipulation, **NumPy** for numerical operations, and **scikit-learn** for preprocessing tools such as *MinMaxScaler*. The SARIMAX notebook implements **statsmodels** to fit seasonal ARIMA models and custom NumPy block-bootstrap for 95% confidence intervals. XGBoost notebook employs **xgboost** and **scikit-learn** regression API to train a gradient-boosting tree on lagged features ( $t-1$ ,  $t-7$ ) and rolling statistics. The LSTM notebooks implement **TensorFlow/Keras** with one serving as a baseline model with two-layer LSTM (64 units per layer, dropout 0.2, Adam optimizer) and the other implements transfer learning on a public perfumes e-commerce series with data-augmentation techniques (Gaussian jitter, time-warping and synthetic bootstrapping).

These three notebooks generate MAE and RMSE on a rolling window train/validation/test split to compare results. Additionally, visual representations of the generated forecasts are available in each notebook for better understanding.

**Table 2: Functional and Non-Functional Requirements**

ID	Type	Requirement
FR-01	Functional	Daily ingestion from <code>combined_final.csv</code> and cleaning (imputation, flags, missing dates).
FR-02	Functional	Produce a 7-day forecast for <code>total_sales</code> with each of the three models.
FR-03	Functional	Compute 95% bootstrap confidence intervals in SARIMAX.
FR-04	Functional	Generate “stock-out risk” alerts when the actual stock + forecast predictions are below the minimum inventory levels established.
PE-01	Performance	Minimum accuracy: $RMSE \leq 4.0$ and $MAE \leq 3.5$ on the hold-out test set for each model.
RE-01	Reliability	Notebooks and pipelines must execute without failures, using fixed seeds to guarantee reproducibility.
MA-01	Maintainability	Code organized into clear sections and documented; version control for notebooks and data; well-structured helper scripts.

#### 4.1 Key Algorithms and Models

- **SARIMAX + Bootstrap:** Selects  $(p,d,q) \times (P,D,Q)$  via AIC/BIC, then applies block-bootstrap on residuals for 95 % confidence bands.
- **XGBoost:** Trains a gradient-boosted tree regressor on lag features ( $t-1, t-7$ ), rolling means, and holiday/campaign dummies, with Bayesian optimization over `learning_rate`, `max_depth`, and `n_estimators`.
- **Baseline LSTM:** Two stacked LSTM layers (64 units each) with dropout 0.2, MSE loss, Adam optimizer (lr  $1e-3$ ).
- **LSTM + Data Augmentation + Transfer Learning:** Pretrained on public perfume time series (50 epochs, batch 32), then fine-tuned on LeBorêt data augmented by Gaussian jitter ( $\sigma=1$  % of range),  $\pm 10$  % time-warp, and synthetic bootstrapping (30 epochs, early stopping).

By isolating each model in a different notebook, we can ensure a clean separation of concerns and easy experimentation. Once the models are evaluated, the best-performing model will be selected and the sales forecast generated will be integrated in COSMOS, the user interface build as part of this project.

## 4.2 Data Structures, Interfaces and Frond-End Components

The `combined_final.csv` schema is listed below:

- Date (DATE)
- total\_sales (INT)
- total\_sessions (INT)
- inventory (INT)
- ignore (BOOL)

The front-end files are divided in three main types:

- HTML files per each interface's page define the structure and container elements for all the charts and tables in the front-end.
- CSS files apply the delimited layout, typography and responsive styling.
- JS files contain the JavaScript logic to fetch CSVs, convert them in arrays, detect alert conditions and initialize Chart.js line and bar charts.

## 4.3 Limitations and Future Extensions

The current limitations include the reliance on a static CSV without live data updates and that only core predictors (sessions, sales and inventory) are used. Additionally, the inventory field in `combined_final.csv` is not fully reliable, during the first months of LeBorêt in 2024, the founder did not input correctly stock levels, which may bias model training and alert thresholds. Although the inventory reliability was considered and mapped for our model training, it is still important to mention the potential limitations of the lack of reliable, complete inventory data.

Future extensions might include a minimal backend API to serve updated CSVs with strong cybersecurity protocols, adding dynamic regressors such as pricing, marketing campaigns and budget. Moreover, a migration to a reactive front-end such as React or Vue may enhance interactivity and integrate advanced forecasting libraries.

## 5 Implementation

This section outlines the implementation of the proposed AI-driven forecasting system for LeBorêt, performed following the Lean Startup methodology. The process involved continuous experimentation, validated learning and iterative delivery of a Minimum Viable Product (MVP) to make sure that the solution directly addressed real operational needs under e-commerce startup's particular conditions.

## 5.1 Lean Architecture Overview

The final system architecture followed the Build-Measure-Learn cycle as the core of the selected methodology (Lean Startup) and it was organized in three modular layers:

- **Data Layer (Build):** Transactional historical data from LeBorêt’s Shopify API, enriched with external series and engineered features. Additionally, key insights from meetings with the founder were also used as a baseline to determine feature selection and interface design priorities.
- **Model Layer (Measure):** This section includes separate forecasting pipelines for SARIMAX + Bootstrap, LSTM variants (2) and XGBoost to test the performance hypothesis.
- **Interface Layer (Learn):** COSMOS interface was iteratively shaped by user feedback to ensure usability and operational fit, plus additional key features with second priority were mapped and illustrated for future work. A 9-page interface was fully developed with real data (when available) and dummy values for future integration.

This modular architecture enables fast prototyping and continuous iteration, which allows changes to one layer without disrupting the other parts of the system.

## 5.2 Data Pipeline for MVP Foundation

The data pipeline marked the foundation of the MVP and was designed to handle early-stage e-commerce data types. The key steps are the following:

- Automated extraction of sales, sessions and inventory data from Shopify.
- Data cleaning and preparation was performed with an additional stock-out tagging verification through direct feedback from the founder.
- Included feature engineering to enable model training with lag variables and calendar effects.
- External data integration from Kaggle for model pretraining (transfer learning) and augmentation techniques.

The outcome was anonymized in a structured dataset “combined\_final.csv”. This unified dataset was used consistently across all model experiments, which enables reproducible, standardized and scalable data flows for future implementations.

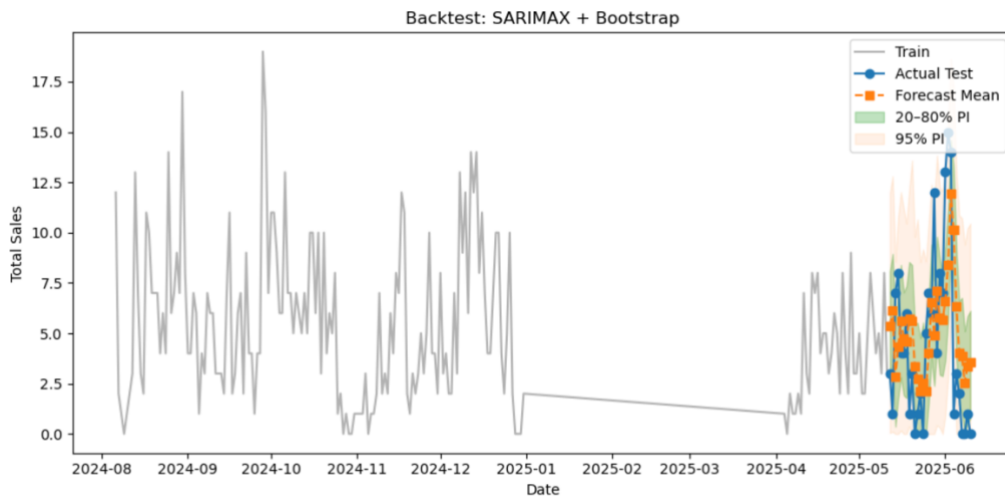
## 5.3 Hypothesis-Driven Forecasting Pipelines

Four model pipelines were built and tested with the project’s specific hypothesis regarding model performance in data constrained situations. Hypotheses:

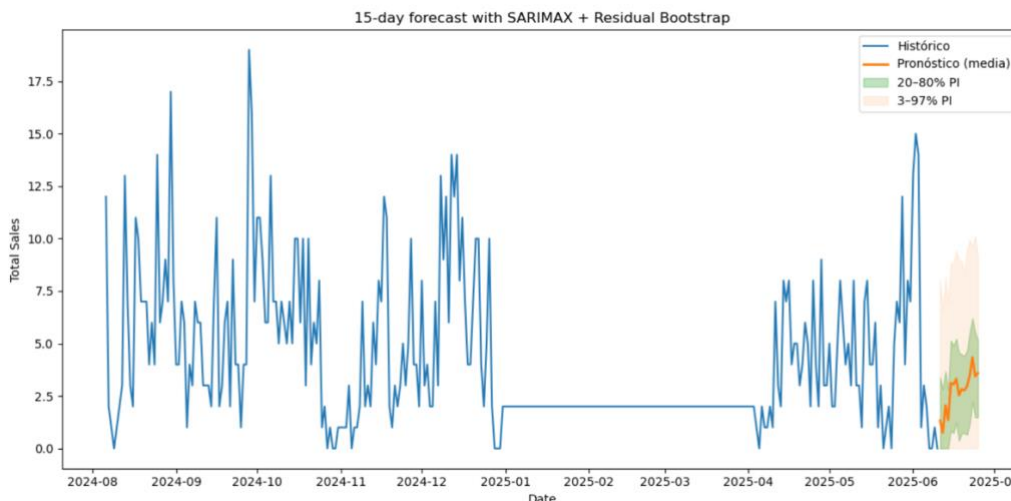
- **SARIMAX – Bootstrap:** statistical models with uncertainty estimation improve decision flexibility in data scarce environments.

- **LSTM Baseline:** Basic sequence models capture better temporal patterns than regressors.
- **LSTM + Transfer Learning + Data Augmentation:** pretrained models augmented with synthetic data generalize better with small datasets.
- **XGBoost:** tree-based models handle better mixed features and outperform deep learning in low-data environments.

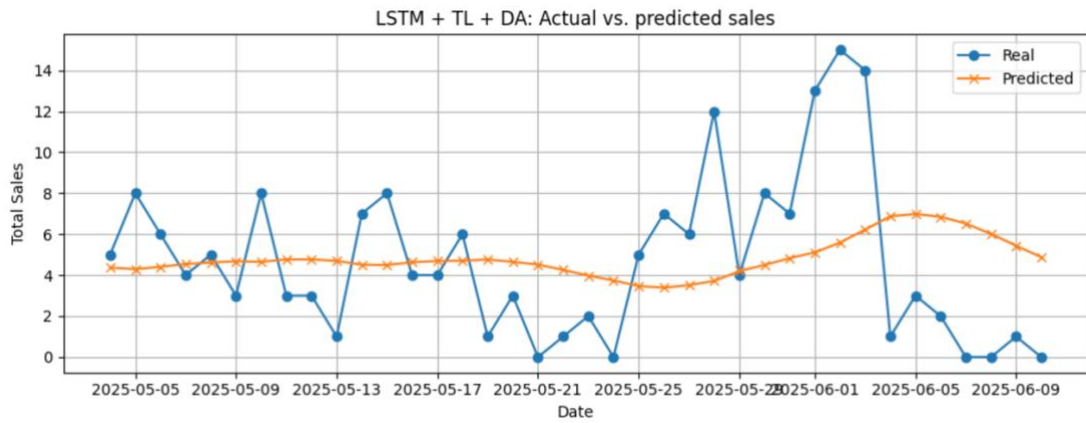
Each model was isolated in a Jupyter notebook to support each experimentation and enabled independent validation. The following figures show each model visual results in each pipeline:



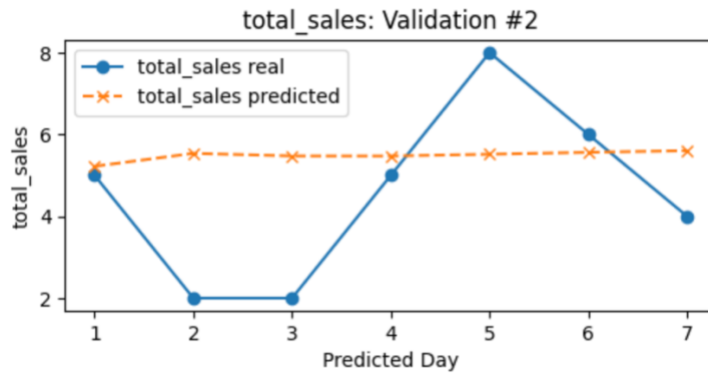
**Figure 1.** SARIMAX + Bootstrap Train/Test Results



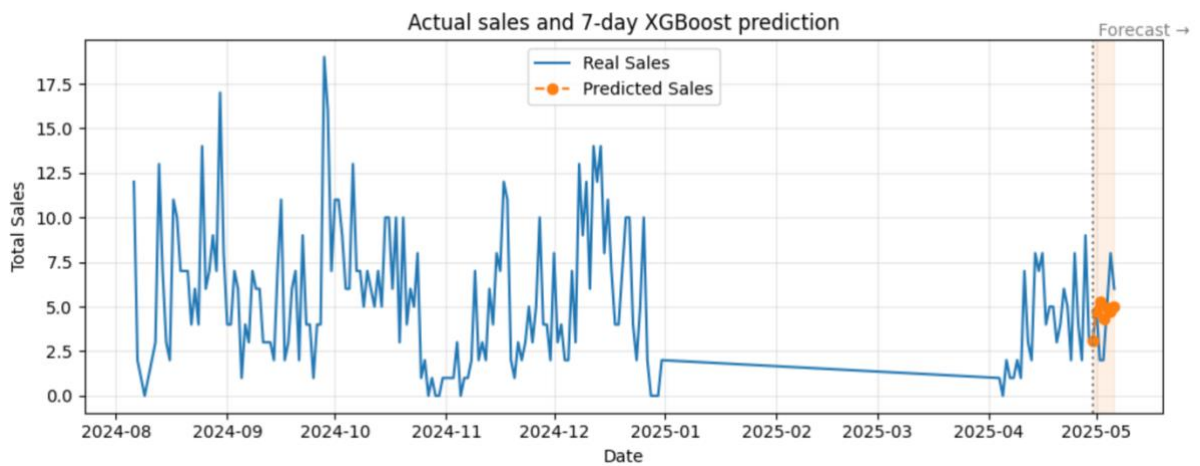
**Figure 2.** SARIMAX + Bootstrap Forecasting Results



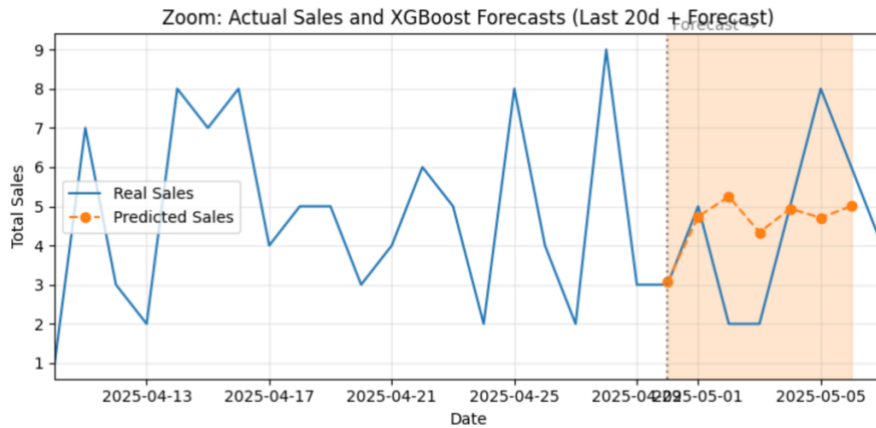
**Figure 3.** LSTM + DA + TL Forecasting Results



**Figure 4.** LSTM Baseline Forecasting Results



**Figure 5.** XGBoost Train/Test Results



**Figure 6.** XGBoost Forecast Results (Zoomed)

### 5.4 MVP User-Centric Design

The MVP was continuously refined and adjusted through usability sessions with LeBorêt’s founder. The following are the prioritized features highlighted by a real startup and were taken as validated learning to design and build the interface:

- **Log-in Page:** Immersive space-themed authentication gateway (*Functional*).
- **Dashboard:** Central command center that displays key performance indications, real-time metrics and system status (*Functional*).
- **Demand Forecasting:** Advanced analytics to predict future sale trends and demand patterns (3 available forecast configurations). Include interactive charts and historical data analysis, conversion rate tracking (*Functional*).
- **Smart Inventory:** Intelligent inventory management system with real-time stock monitoring, automated reorder point calculations and critical stock alerts (*Functional*).
- **Fulfilment:** Order processing and logistics optimization to track shipments and managing workflows (*Functional*).
- **Alarms and Alerts:** System notification center for monitoring critical events, inventory alerts, and operational warnings (*Functional*).
- **To Do Board:** Kanban style management for organizing all the operational tasks across all modules (*Functional*).
- **Calendar:** Scheduling interface for planning and tracking tasks, events and operational activities (*Functional*).
- **Partners:** Supplier and business partner management to display contact types, information and partnership analytics (*Functional*).
- **Document Base:** Centralized document repository (*Future work*).

**Note:** the features labelled as “Functional” are fully implemented in the first MVP. Features marked as “Future Work” are visible in the interface but not yet operational and are planned in the scope of the second MVP. The link for full COSMOS visualization is [here](#).

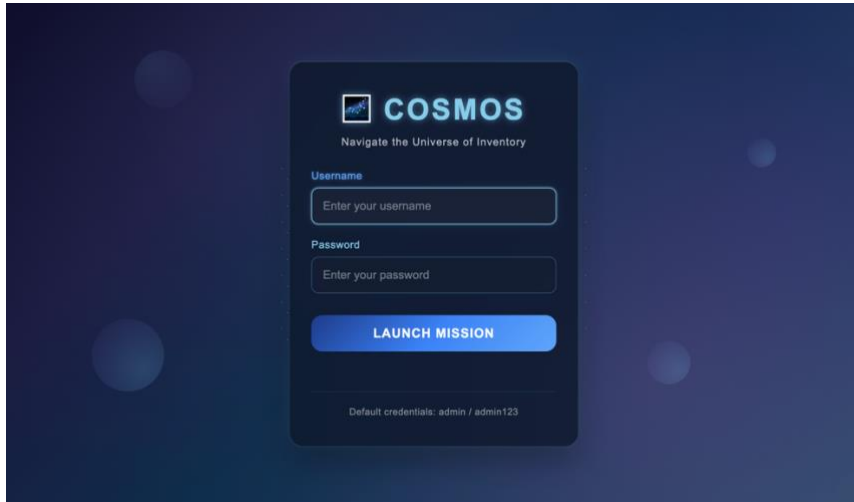


Figure 7. Log in page visualization



Figure 8. Demand Forecasting main graph visualization



Figure 9. Smart Inventory Page

Order ID	Product	Zone	Days Since Order	Priority Score	Assigned To	Status	Actions
ORD-0013	Tundra Eau de Parfum	Zona Sur	6	75	German Gomez	READY	Start, Assign, View
ORD-0041	Set LeBoret No. 1	Zona Sur	5	70	German Gomez	READY	Start, Assign, View
ORD-0042	Set Descubriendo LeBoret	Zona Centro	6	70	German Gomez	PENDING	Start, Assign, View
ORD-0001	Sahara Eau de Parfum	Zona Centro	6	65	Valeria Elizondo	PROCESSING	Start, Assign, View
ORD-0005	Tundra Eau de Parfum	Zona Sur	5	65	German Gomez	PROCESSING	Start, Assign, View
ORD-0027	Tundra Eau de Parfum	Zona Centro	6	65	German Gomez	READY	Start, Assign, View

Figure 10. Fulfilment Assignment and Monitoring Orders Table

**Alarms & Alerts**  
Monitor critical events and receive intelligent notifications to keep your operations running smoothly.

- Critical Alerts** (3): Require immediate attention.
- Warning Alerts** (7): Action needed soon.
- All Systems** (4): Operating normally.

**Current Alerts**

- Stock Out - Set LeBoret No. 1**: Only 63 units remaining. Reorder point reached for SKU #1. (Smart Inventory, 5 min ago)
- AromaTech SA Supplier Delay**: 15-day lead time exceeded. New delivery estimate: 3 days overdue. (Fulfillment, 1 hour ago)
- Forecast Accuracy Drop - Manhattan EDP**: Prediction accuracy fell to 65% (below 75% threshold). Model needs retraining. (Demand Forecasting, 2 hours ago)

Figure 11. Alarms and Alerts Dashboard

**Mission Control Board**  
Centralized command center for all operational tasks and alerts.

Navigation: All Modules, Demand Forecasting, Fulfillment, Smart Inventory, Alarms & Alerts

**To Do**

- Update demand forecast model** (HIGH): Incorporate latest sales data into Q3 forecasting model. Due: 2025-08-10.
- Model accuracy validation** (HIGH): Compare predicted vs actual sales for July 2025. Due: 2025-08-08.
- Critical inventory rebalancing** (HIGH): Rebalance stock levels for products below safety threshold.

**In Progress**

- Seasonal trend analysis** (MEDIUM): Analyze Q4 holiday patterns for 2025 forecasting. Due: 2025-08-15.
- Update reorder points** (MEDIUM): Recalculate reorder points based on Q3 demand patterns. Due: 2025-08-12.
- ABC classification update** (HIGH): Reclassify inventory based on value and velocity analysis.

**Completed**

- Weekly forecast analysis** (LOW): Generate automated weekly demand analysis report. Due: 2025-08-05.
- Holiday impact assessment** (MEDIUM): Evaluate Independence Day sales impact on forecast model. Due: 2025-08-07.
- Vendor performance review** (LOW): Assess AromaTech delivery times and adjust lead time calculations.

Figure 12. To-Do Board with Classification

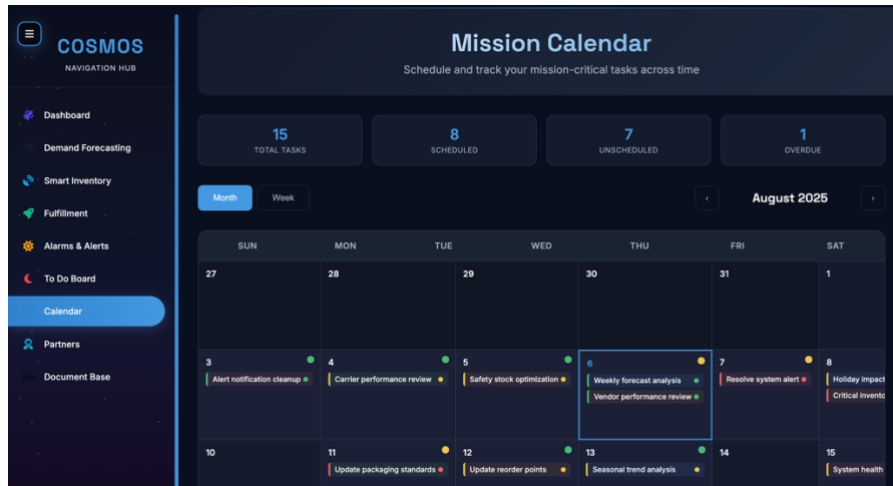


Figure 13. Calendar and Today's Activities Dashboard

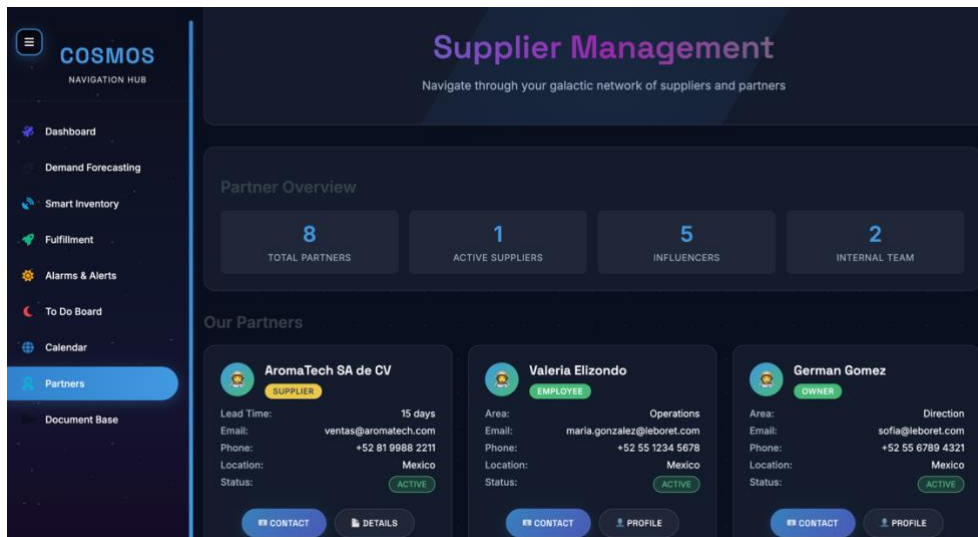


Figure 14. Partner Management Board

The front-end was implemented using lightweight and cost-effective web technologies including HTML, CSS and JS, with Chart.js and CSV-driven data binding, which enabled fast iteration while ensuring data protection by avoiding direct use of LeBorêt's Shopify API in the first MVP.

### 5.5 Lean Startup Advantages in Implementation

The solution architecture based on the Lean Startup approach ensured that the features were developed on real user feedback and actual business needs rather than assumptions or theoretical concepts. The key outcomes include feature selection with validated learning, early deployment of usable components to accelerate feedback and modular pipelines and interface that allowed parallel experimentation and refinement.

In summary, this implementation delivered a functional first MVP, COSMOS, aligned with technical objectives and real-world usability in relatively unexplored domain, startup

ecosystems. It contributes to the democratization of AI while offering a practical and scalable solution to forecast demand and efficient inventory management tailored to the realities of early-stage startups.

## **6 Evaluation**

The evaluation performed in this project aims to assess three core points: the predictive performance of the selected forecasting models under limited-data conditions (1), practical utility of the proposed MVP COSMOS interface in a real e-commerce startup context (2) and get user validation through direct feedback sessions (3). These evaluation elements were selected to reflect technical success, alignment with real operational needs and coordination with the Lean Startup framework.

### **6.1 Evaluation Methods**

The evaluation measures the quantitative and qualitative performance of the proposed solution. Quantitatively, model performance was assessed using RMSE and MAE on the selected dataset to predict. The SARIMAX model specifically was also evaluated based on 95% prediction interval coverage. Additionally, a rolling-window time series validation method was applied to mimic real-world deployment conditions to assess results. Qualitatively, the MVP was tested in two structured usability sessions with LeBorêt's founder focusing on alignment with operational workflows and needs, clarity and ease of use.

### **6.2 Results Summary**

The four evaluated models showed different strengths. SARIMAX + Bootstrap achieved an RSME of 3.78 and a MAE of 3.21, with a 93.33% coverage of its 95% prediction intervals. The LSTM Baseline model achieved an RMSE of 3.89 and MAE of 3.12. The Enhanced LSTM with transfer learning and data augmentation techniques reached an RMSE of 4.11 and a MAE of 3.37. Finally, XGBoost presented a performance range depending on the selected forecast horizon, from an RMSE of 3.22 to 4.94 and a MAE of 2.61 to 4.10. While the XGBoost and LSTM models offered a more competitive accuracy, SARIMAX provided a greater interpretability and flexibility through its prediction intervals.

**Table 3.** Comparative Results of the Models

Model	MAE	RMSE	95% Coverage	Comments
SARIMAX + Bootstrap	3.21	3.78	93.33%	High coverage and easy interpretation, ideal for flexible planning with limited resources. Less accurate in highly volatile contexts.
LSTM + TL + DA	3.37	4.11	N/A	Strong generalization in data-scarce situations and adapts well to irregular patterns. High complexity and computational costs.
LSTM Baseline	3.12	3.89	N/A	Good performance with simple model architecture. Tends to underestimate real demand peaks.
XGBoost	2.61-4.10	3.22-4.94	N/A	Efficient and robust performs well in small datasets. Lacks long-term temporal modelling and it doesn't provide uncertainty estimations.

While machine learning and deep learning models showed competitive performance in prediction accuracy, SARIMAX with Bootstrap stands out as the most suitable model for startups like LeBorêt. This is because it has the ability to generate well-calibrated confidence intervals with 93.33% coverage, which is particularly valuable for inventory decision-making under fluctuating and budget constrained environments. Prediction intervals give startups flexibility to plan using conservative, neural and optimistic demand scenarios, which is crucial to have the opportunity to adjust inventory procurement to cash flow availability. This feature, which enables flexibility, is vital in early-stage startups, where forecasting errors can directly impact their survival and growth. In short, despite some models showed slightly lower average errors, SARIMAX + Bootstrap offers a balanced combination between accuracy levels, interpretability and operational value, making it the best choice for this particular case study.

The project met the predefined design specification performance PE-01, with the models evaluated achieving and  $RMSE \leq 4.0$  and  $MAE \leq 3.5$ .

### 6.3 MVP Evaluation

COSMOS was developed following the Lean Startup Methodology, which consists in iterative Build-Measure-Learn cycles. Feature and page selection was based on a qualitative/quantitative assessment conducted with Leborêt's founder, focusing on three key criterias: usability, user experience (UX), and operational value (OV). The results are presented in the following table and evaluated in a 1 to 10 range:

**Table 4.** Combined quantitative + qualitative evaluation table for COSMOS

Page	Usability	UX	OV	Justification
<b>Log-in</b>	9	10	7	Smooth authentication flow with engaging space-theme. Quick to load.
<b>Dashboard</b>	9	8	10	Centralized KPIs, high valuable for decision-making. Learning curve for new users.
<b>Demand Forecasting</b>	9	9	10	Strong analytics with multiple configurations, usability slightly impacted by complexity.
<b>Smart Inventory</b>	9	8	10	Stock tracking and automated reordering points, high operational benefit. UX could be more visual, too much information.
<b>Fulfilment</b>	9	8	9	Solid order tracking and workflow tools.
<b>Alarms &amp; Alerts</b>	9	9	10	Fast and visible notification for critical events, intuitive and very impactful in operations.
<b>To Do Board</b>	10	9	8	Clear Kanban view, operational value relies on user adoption.
<b>Calendar</b>	8	9	9	Easy scheduling with good visibility, easy to follow and understand.
<b>Partners</b>	8	8	9	Organized contact and partnership data, depends on data completeness.
<b>Document Base (Future)</b>	-	-	-	Planned centralized repository that will support management once implemented.

COSMOS obtained an overall 8.8/10 score, with the Demand Forecasting page ranking highest at 9.3/10, followed by Dashboard (9.0), Smart Inventory (9.0) and To Do Board (9.0). The high operational value demonstrated for Leborêt strongly suggests that the interface will be equally useful and effective for similar startups.

#### 6.4 Limitations and Lessons Learned

The main limitations in this project included that the training dataset contained missing values and noise, which are common in startup environments. Also, the selected pretrained data used for the enhanced LSTM may not represent fully the niche behaviour of LeBorêt's products, which can affect results. The key lessons learned are listed below:

1. Confidence intervals add a significant value for planning flexibility under uncertainty and limited-data scenarios.
2. Forecasting accuracy alone is insufficient in this case; business context and interpretability are key additional factors to consider.
3. Continuous and structured user feedback leads to tailored and highly usable features and higher adoption of AI-driven solutions.
4. Modular implementation is key in resource-constrained environments and support rapid iteration and future applications.

## 7 Conclusion and Future Work

This project addressed the research question: *How can an AI-powered decision-support system be designed to deliver reliable demand forecasts and operational recommendations for startups with limited data availability?* Using LeBorêt, a real Mexican ecommerce startup, as a case study, the primary objectives were comparing AI-driven models and select the best one for startup environment and to build a scalable, functional MVP capable of integrating the selected forecasting model with operational modules to enhance decision-making efficiency and centralized control of key operations.

From the model comparison, a quantitative evaluation using MAE and RMSE was implemented to select the one showing strong adaptability to non-linear patterns, typical in startup contexts. SARIMAX + Bootstrap was ultimately selected as the default model in COSMOS for operational deployment for three main reasons:

1. Provided the most stable and interpretable results, which aligns with the operational decision cycles of LeBorêt startup.
2. It requires few computational resources compared to the other models, which ensures faster response times and low budget adaptability.
3. The statistical structure of this model makes it more transparent for non-technical stakeholders, which facilitates model's adoption and trust.

The COSMOS platform achieved an overall 8.8/10 score measuring usability, user experience and operational value. These high scores demonstrate that this interface not only meets the project's functional goals, but it also provides clear operational value for the client. This success is mainly driven by the Lean Startup Methodology followed in the entire development of the interface, which let only features and pages approved and validated from real users to be selected.

Overall, the project did achieve its objective: it validated that AI-driven forecasting with multiple configurable models is feasible for startups with scarce data, provides a scalable architecture and offer actionable insights that can be extended for other similar early-stage startups. Future work for further refinement of the proposed solution includes a live integration that connect the interface directly to the startup's data APIs, such as Shopify for real time forecasting, automated inventory recommendations with optimization algorithms that can generate actionable order quantities, expanded model library to explore other models such as hybrid and/or transformer-based for further performance improvements. Finally, a broader testing is suggested to fully deploy the system in other startups to assess scalability.

## References

- Abuzaid, A. N., & Alsou, M. K. K. (2024, April). AI and entrepreneurship: Enablers, obstacles, and startups' role in shaping the future economy. In *2024 International Conference on Knowledge Engineering and Communication Systems (ICKECS)* (Vol. 1, pp. 1–6). IEEE.
- Agarwal, K., Dheekollu, L., Dhama, G., Arora, A., Asthana, S., & Bhowmik, T. (2021). Deep learning-based time series forecasting. In *Deep Learning Applications, Volume 3* (pp. 151–169). Springer Singapore.
- Carayannis, E. G., Dumitrescu, R., Falkowski, T., & Zota, N. R. (2024). Empowering SMEs—Harnessing the potential of Gen AI for resilience and competitiveness. *IEEE Transactions on Engineering Management*.
- Chang, K. C., & Huang, W. K. (2024, July). Design of retail product recognition by transfer learning and fine-tuning techniques. In *2024 International Conference on Consumer Electronics-Taiwan (ICCE-Taiwan)* (pp. 449–450). IEEE.
- Hirt, R., Kühn, N., Peker, Y., & Satzger, G. (2020, June). How to learn from others: Transfer machine learning with additive regression models to improve sales forecasting. In *2020 IEEE 22nd Conference on Business Informatics (CBI)* (Vol. 1, pp. 20–29). IEEE.
- Jessica, & Suryasari. (2024, December). An ensemble approach combining LSTM and random forest for forecasting product stock expenditure in retail. In *2024 International Conference on Information Technology Systems and Innovation (ICITSI)* (pp. 262–269). IEEE.
- Jia, J., & Yan, R. (2023, December). Research on short-term forecast of product demand based on SARIMAX model. In *2023 International Conference on Applied Physics and Computing (ICAPC)* (pp. 449–454). IEEE.
- Lawton, R., Boswell, S., & Smice, K. C. (2023, December). The GM AI Foundry: A model for upskilling SMEs in responsible AI. In *2023 IEEE Symposium Series on Computational Intelligence (SSCI)* (pp. 1781–1787). IEEE.
- Li, Q., Qiu, Y., Liu, W., & Luan, J. (2024, August). Improving load forecasting accuracy through data augmentation and TCNs-transformer. In *2024 4th Power System and Green Energy Conference (PSGEC)* (pp. 65–69). IEEE.
- Lv, J. (2024, February). Research on inventory management and demand forecasting of e-commerce platform based on ARIMA and LSTM models. In *2024 IEEE 3rd International Conference on Electrical Engineering, Big Data and Algorithms (EEBDA)* (pp. 494–499). IEEE.
- Manos, A., Filippidou, D. E., Pavlidis, N., Karanasios, G., Vachtanidis, G., Mallidis, I., ... & Ayfantopoulou, G. (2024, December). Dynamic demand forecasting with AI: Maximizing supply chain value. In *2024 International Conference on Engineering and Emerging Technologies (ICEET)* (pp. 1–4). IEEE.
- Mao, S., & Sejdić, E. (2022). A review of recurrent neural network-based methods in computational physiology. *IEEE Transactions on Neural Networks and Learning Systems*, 34(10), 6983–7003.
- Özen, S., Atalay, V., & Yazici, A. (2019, November). Comparison of predictive models for forecasting time-series data. In *Proceedings of the 3rd International Conference on Big Data Research* (pp. 172–176).
- Rajendran, V., Tiwari, A., & Tripathi, V. (2024, January). Application of AI in determining the strategies for the startups. In *2024 International Conference on Healthcare Innovations, Software and Engineering Technologies (HISSET)* (pp. 374–376). IEEE.
- Ries, E. (2011). *The lean startup: How today's entrepreneurs use continuous innovation to create radically successful businesses*. Crown Currency.
- Sharma, A., Sahu, B., Kishore, M., Kumari, K., & Shubhnath, S. S. (2024, October). Implementation of models for demand forecasting for e-commerce using time series forecasting. In *2024 IEEE 1st International Conference on Green Industrial Electronics and Sustainable Technologies (GIEST)* (pp. 1–6). IEEE.
- Sharma, T., Prasad, S. K., Prasad, S., Verma, I., & Sharma, A. (2024, December). Forecasting stock market volatility using XGBoost: A time series analysis. In *2024 International Conference on Communication, Control, and Intelligent Systems (CCIS)* (pp. 1–6). IEEE.
- Sheng, F., & Jia, L. (2020, September). Short-term load forecasting based on SARIMAX-LSTM. In *2020 5th International Conference on Power and Renewable Energy (ICPRE)* (pp. 90–94). IEEE.
- Siami-Namini, S., Tavakoli, N., & Namin, A. S. (2018, December). A comparison of ARIMA and LSTM in forecasting time series. In *2018 17th IEEE International Conference on Machine Learning and Applications (ICMLA)* (pp. 1394–1401). IEEE.
- Singh, J., Shelke, N. A., Upreti, K., Saiyad, F. B. J., Divakaran, P., & Madhukar, K. S. (2024, July). Deep learning advancements in e-commerce supply chain management in forecasting and optimization strategies. In *2024 IEEE 4th International Conference on Sustainable Energy and Future Electric Transportation (SEFET)* (pp. 1–6). IEEE.
- Veretennikova, N., & Vaskiv, R. (2018, September). Application of the lean startup methodology in project management at launching new innovative products. In *2018 IEEE 13th International Scientific and*

- Technical Conference on Computer Sciences and Information Technologies (CSIT)* (Vol. 2, pp. 169–172). IEEE.
- Wang, Y., & Mi, X. (2020, May). A comparative study on demand forecast of car sharing users based on ARIMA and LSTM. In *2020 5th International Conference on Electromechanical Control Technology and Transportation (ICECTT)* (pp. 565–574). IEEE.
- Yang, X., Zhang, Z., Cui, X., & Cui, R. (2021, May). A time series data augmentation method based on dynamic time warping. In *2021 International Conference on Computer Communication and Artificial Intelligence (CCAI)* (pp. 116–120). IEEE.
- Zhao, Y. (2024). Research on e-commerce retail demand forecasting based on SARIMA model and K-means clustering algorithm. *Academic Journal of Science and Technology*, *10*(3), 226–231.

# Appendix 1 – AI Cover Page

## AI Acknowledgement Supplement

### Practicum

#### Application of AI in E-Commerce for Startups with Limited Data: A Lean Startup Approach to Demand Forecasting and Inventory Optimization

Your Name/Student Number	Course	Date
Ana Paula Elizondo de la Garza / 23140879	Practicum	August 10, 2025

This section is a supplement to the main assignment, to be used if AI was used in any capacity in the creation of your assignment; if you have queries about how to do this, please contact your lecturer. For an example of how to fill these sections out, please click [here](#).

### AI Acknowledgment

This section acknowledges the AI tools that were utilized in the process of completing this assignment.

Tool Name	Brief Description	Link to tool
ChatGPT	Idea generation for layouts, word synonyms and spelling checks.	<a href="https://chat.openai.com/">https://chat.openai.com/</a>

### Description of AI Usage

This section provides a more detailed description of how the AI tools were used in the assignment. It includes information about the prompts given to the AI tool, the responses received, and how these responses were utilized or modified in the assignment. **One table should be used for each tool used.**

[Insert Tool Name]	
ChatGPT was used to provide ideas for layouts, word synonyms and spelling checks in the project.	
Can you help me remember the requirements for the Methodology section of my report form the official Handbook?	<p>Yes — the <i>official</i> requirements for the <b>Research Methodology</b> section are in Section 5.4.3 of the MSc Practicum/Research Practicum/Internship Handbook, and they line up with the Research Project Report Template.</p> <p>From the <b>Handbook</b>: The Research Methodology section should clearly detail:</p> <ul style="list-style-type: none"> <li>• <b>The steps followed</b> in the research process (the “process” itself).</li> <li>• <b>The data required</b> to answer your research question(s), e.g., database records, tools, software, algorithms, metrics, etc.</li> <li>• <b>How you obtained and used the data</b> in your project.</li> <li>• <b>How the data and results</b> collected or measured were analysed.</li> </ul> <p>It should also connect your data requirements to <b>specific research objectives</b> — essentially creating a “shopping list” of the data you need to collect to meet those objectives.</p> <p>From the <b>Report Template</b>: Your methodology section should:</p> <ul style="list-style-type: none"> <li>• Give a <b>completely accurate description</b> of the research procedure you followed.</li> <li>• Include details of <b>equipment used, techniques applied, and scenario/case study setup</b>.</li> <li>• Explain <b>how raw data was gathered/compiled and analysed</b>.</li> <li>• Describe <b>statistical techniques</b> used on the data.</li> <li>• Detail <b>all steps</b> from data collection to final results.</li> <li>• Justify your choices using <b>appropriate literature</b>, ideally linking to your related work review.</li> </ul> <p>Both sources stress that this section is crucial for examiners to confirm that you followed a <b>rigorous, appropriate scientific process</b> and that your decisions are clearly rationalised.</p>
can you help me find synonyms for the word "Nevertheless"	<p>Sure — here are some good synonyms for "<b>Nevertheless</b>", depending on the tone you want:</p> <ul style="list-style-type: none"> <li>• <b>However</b></li> <li>• <b>Nonetheless</b></li> <li>• <b>Even so</b></li> <li>• <b>All the same</b></li> <li>• <b>Yet</b></li> </ul>

- **Still**
- **Be that as it may**
- **In spite of that**
- **Though** (at the start or end of a sentence for a softer feel)
- **Regardless**

If you'd like, I can also group them by **formal**, **neutral**, and **conversational** tone so you can pick the best fit for your methodology section. Would you like me to do that?

## Evidence of AI Usage

This section includes evidence of significant prompts and responses used or generated through the AI tool. It should provide a clear understanding of the extent to which the AI tool was used in the assignment. Evidence may be attached via screenshots or text.

### Additional Evidence:

HandbookGPT 5 ▾

↑ Comp

Can you help me remember the requirements for the Methodology section of my report form the oficial Handbook?

Yes — the *official* requirements for the **Research Methodology** section are in Section 5.4.3 of the MSc Practicum/Research Practicum/Internship Handbook, and they line up with the Research Project Report Template.

From the **Handbook**:

The Research Methodology section should clearly detail:

- **The steps followed** in the research process (the "process" itself).
- **The data required** to answer your research question(s), e.g., database records, tools, software, algorithms, metrics, etc.
- **How you obtained and used the data** in your project.
- **How the data and results** collected or measured were analysed.

It should also connect your data requirements to **specific research objectives** — essentially creating a "shopping list" of the data you need to collect to meet those objectives.

From the **Report Template**:

Your methodology section should:



+ Pregunta lo que quieras



### Additional Evidence:

HandbookGPT 5 ▾

↑ Compartir

can you help me find synonyms for the word "Nevertheless"

Sure — here are some good synonyms for "**Nevertheless**", depending on the tone you want:

- **However**
- **Nonetheless**
- **Even so**
- **All the same**
- **Yet**
- **Still**
- **Be that as it may**
- **In spite of that**
- **Though** (at the start or end of a sentence for a softer feel)
- **Regardless**

If you'd like, I can also group them by **formal**, **neutral**, and **conversational** tone so you can pick the best fit for your methodology section. Would you like me to do that?

