SME Sales Forecasting and Drift Detection Using Neural Networks and Public Macroeconomic Indicators

Kyaw May Pyone Shinn

x24123455

MSc in Data Analytics

National College of Ireland

Abstract

Small and medium-sized enterprises (SMEs) are highly exposed to macroeconomic volatility but often lack advanced analytics to anticipate adverse performance trends. This research investigates whether combining time-series forecasting with drift detection can provide early warning signals for SME sales instability. Public macroeconomic indicators from Alpha Vantage and FRED were integrated with a synthetically generated panel of 50 SME-like firms observed daily from 2021 to 2023. After extensive feature engineering, a multiple linear regression (MLR) model and a Long Short-Term Memory (LSTM) network were trained to forecast next-day sales, while a feedforward autoencoder was used to detect deviations from normal operational patterns. The Multiple Linear Regression model achieved an R^2 of 0.8547 and an RMSE of 180.1445 on the held-out test period. The LSTM achieved an R^2 of 0.8322 and an RMSE of 198.9795 on the same test period. Compared with the baseline regression model, the LSTM's RMSE was 10.45% higher, and its R^2 was 2.26% lower, indicating that the deep learning model did not provide a measurable performance gain under SME-style data constraints. The autoencoder attained high recall (0.60) at the cost of low precision (0.02), making it suitable as a sensitive drift indicator. The study contributes a reproducible pipeline that integrates macroeconomic data, forecasting models and anomaly detection into a business intelligence-oriented framework, and provides evidence that well-engineered linear models can match or exceed deep learning performance in SME-scale settings.

Keywords: Business Intelligence, LSTM, Autoencoder, Anomaly Detection, Time-Series Forecasting

1. INTRODUCTION

Small and medium enterprises (SMEs) operate in environments where market conditions, input costs, and customer demand can change rapidly. These organisations often have limited access to advanced analytical tools that can forecast future performance or identify early signs of operational instability. Larger firms typically use predictive systems and automated monitoring, but SMEs rely on historical reports that do not support forward-looking decisions. This creates a gap in analytical capability. The need for more accessible forecasting and anomaly detection methods for SMEs is therefore an important motivation for this study.

The motivation is strengthened by the lack of research that combines short-term forecasting and drift detection in a single, reproducible workflow. Many commercial business intelligence tools focus on descriptive visualisation rather than predictive analysis. Academic studies usually examine forecasting or anomaly detection separately and often use proprietary datasets that are not available for replication. Real SME data is difficult to obtain due to confidentiality concerns, which restricts experimentation. This study addresses these challenges by investigating whether neural network models can work together with synthetic SME data and public macroeconomic indicators to support predictive insights for SME decision making.

The research problem arises from this gap. It concerns the need to understand whether forecasting models and anomaly detection models can be applied to SME KPI data in a way that is reliable, interpretable, and aligned with the requirements of a lightweight business intelligence workflow. The study therefore examines the forecasting performance of the models, their behaviour under changes in macroeconomic conditions, and their ability to detect unusual patterns or drift.

This leads to the research question guiding the study: **How can neural network models (LSTM and Autoencoder), supported by publicly available macroeconomic indicators, be used to forecast SME KPIs and detect performance drift within a reproducible analytics pipeline suitable for business intelligence integration**? This question highlights both forecasting and drift detection and frames the research as an evaluation of feasibility. To answer this question, the research pursues four objectives: to analyse existing approaches to time-series forecasting and drift detection in related domains; to construct a synthetic but realistic SME dataset aligned with publicly available macroeconomic indicators; to develop and evaluate forecasting models and an anomaly-detection method suited to SME-like data constraints; and to assess how these components can be integrated into a decision-support-oriented analytics pipeline. These objectives emphasise investigation, analysis and evaluation rather than system building, as required for academic research.

The contribution of this study is twofold. First, it provides empirical evidence on the relative performance of linear and deep learning forecasting models in SME-scale settings, demonstrating that feature-rich linear models can rival or exceed the accuracy of more complex architectures. Second, it evaluates the practicality of autoencoder-based reconstruction error as a drift indicator in SME contexts, showing how such models behave under macroeconomic stress and firm-level variability. Combined, these findings contribute to the understanding of predictive analytics for SMEs and propose a reproducible framework suitable for further academic and practitioner development.

The remainder of this report is organized into eight sections. Section 2 presents a critical literature review covering forecasting methods, reconstruction-based drift detection, synthetic data generation, and business intelligence adoption in SMEs. Section 3 outlines the research methodology, detailing the overall methodological approach, data requirements, ethical considerations, and the evaluation strategy. Section 4 describes the system architecture and implementation, including the modelling pipeline, autoencoder drift-detection component, and dashboard integration. Section 5 documents the data preparation procedures, including data acquisition, synthetic SME dataset construction, preprocessing steps, and feature engineering. Section 6 presents the experimental results and evaluates the performance of the forecasting and drift-detection models. Section 7 provides a detailed discussion of the findings, examining their

implications and positioning them within the broader academic context. Section 8 concludes the study by summarizing the contributions, acknowledging limitations, and proposing directions for future research.

2. LITERATURE REVIEW

This literature review critically evaluates prior work in time-series forecasting, anomaly detection, external signal integration, synthetic data generation, and business intelligence (BI) adoption in SMEs. It compares methodological approaches, identifies their limitations, and clarifies how the existing evidence informs and motivates the present study.

2.1. Time-Series Forecasting Models

Time-series forecasting is well studied, and traditional statistical models such as ARIMA and exponential smoothing have been widely applied across economic and financial domains. (Siami-Namini, Tavakoli and Namin, 2019) conducted an influential comparison between ARIMA and deep learning models, including LSTM and GRU networks. Using financial time-series data, they found that LSTM achieved lower error values (RMSE ≈ 0.032) compared to ARIMA, demonstrating the strength of recurrent neural architectures in capturing non-linear temporal relationships. Their conclusion, that LSTMs outperform classical models, has influenced much of the recent deep learning literature. However, their data consisted of long, high-frequency financial series with strong temporal depth. These data conditions are very different from SME environments, where historical records are shorter and often noisier. As such, while their findings demonstrate the potential of LSTMs, the direct transferability to SME KPI forecasting remains uncertain.

Similarly, (Lim, Arik, Loeff and Pfister, 2020) reinforced the advantages of deep learning architectures by surveying a broad family of models, including Transformers, LSTMs and temporal convolutional networks. They emphasized that deep models excel when long-term dependencies, non-linear interactions and exogenous variables are present. Yet, they also noted that these models require considerable data, computational resources and tuning characteristics misaligned with many SME settings. This strengthens the motivation for evaluating the true marginal benefit of LSTMs when operating under SME-like constraints.

Another relevant study is (Cao, Li and Li, 2019), who examined multivariate forecasting using LSTM with macroeconomic variables. Their results showed that incorporating external signals such as inflation or interest rate indicators into LSTM forecasting improved financial time series accuracy. Their study supports the central premise of this report, that external indicators matter. However, their evaluation was limited to stable economic periods and did not examine performance during macroeconomic shocks. This is an important consideration for SME analytics, where macro shocks often drive volatility in operational KPIs.

The literature demonstrates that LSTMs can outperform traditional models under favorable data conditions, but their superiority is not guaranteed in constrained environments such as SMEs. Most existing work is based on financial markets rather than firm-level KPIs. This gap justifies comparing a simple linear model and an LSTM within an SMEstyle dataset that includes macroeconomic drivers, shorter histories and operational irregularities.

2.2. Forecasting Models for Business KPIs

Several studies have explored forecasting techniques for business KPIs. Traditional statistical models like ARIMA have long been used for their simplicity. However, recent research shows that deep learning methods are more effective in capturing complex patterns in financial data. (Siami-Namini, Tavakoli and Namin, 2019) compared ARIMA, LSTM, and BiLSTM, concluding that BiLSTM achieved the highest accuracy in stock market forecasts. Similarly, (Yamak, Yujian and Gadosey, 2019) found that GRU outperformed both ARIMA and LSTM in cryptocurrency prediction when the dataset had nonlinear characteristics. These studies collectively suggest that model performance depends strongly on the specific characteristics of the dataset and volatility of the domain.

(Cao, Li and Li, 2019) introduced a hybrid model combining CEEMDAN and LSTM, which improved prediction accuracy by decomposing noisy financial data but was able to achieve lower MAE than standard models. This decomposition strategy is effective in highly volatile contexts but increases computational complexity, a drawback for SME deployment. (Pietukhov, Ahtamad, Faraji-Niri and El-Said, 2023) developed a neural network-based model for binary KPI forecasting in supply chains and reported improved performance over logistic regression alone but the dataset involved binary outcomes rather than continuous sales metrics. (Ahamed, Vijayasankar, Thenmozhi, Rajendar, Bindu and Subha Mastan Rao, 2023) tested LSTM, GRU, and CNN in operational forecasting and found LSTM to be the most accurate, though computationally intensive.

Although deep learning models consistently show higher accuracy than ARIMA, most studies were conducted offline or in controlled experiments. Their integration into live BI dashboards remains limited. This motivates the current research, which applies LSTM forecasting to real-time SME dashboards using public APIs.

Focusing on LSTM is a practical first step for SMEs because it learns both short and long temporal dependencies that appear in KPIs with seasonality, promotions, and shocks. Comparative evidence shows that deep recurrent models often outperform linear baselines on nonlinear financial and operational series, and this supports the use of LSTM as a first production model in a BI dashboard that must be reliable and explainable for managers. It is still useful to position LSTM against strong modern alternatives, so expectations are clear.

Existing KPI forecasting research shows strong performance from deep learning but lacks evaluation on SME-style datasets that combine internal KPIs, external macroeconomic factors and short time horizons. This supports the need to test whether LSTM meaningfully improves predictions compared with a simpler regression model under SME constraints, one of the key motivations of this report.

2.3. Anomaly and Drift Detection Techniques

Anomaly detection in time-series data is essential for identifying irregular patterns, structural breaks, or performance drift. Autoencoder-based approaches have become a common technique for unsupervised anomaly detection due to their ability to learn compressed representations of normal behaviour. (Audibert, Michiardi, Guyard, Marti and Zuluaga, 2020) demonstrated that deep autoencoders outperform classical approaches such as PCA or one-class SVM in detecting anomalies in high-dimensional data, achieving an AUC above 0.90 in benchmark datasets. Although the

results are promising, their experiments focused on static anomaly detection in benchmark datasets and did not examine slow drift or macro-driven deviations, issues highly relevant to SMEs.

(Malhotra, Ramakrishnan, Anand, Vig, Agarwal and Shroff, 2016; Pavan Kumar and Ravi Teja, 2025) introduced LSTM-based autoencoders for sequence reconstruction and reported strong performance in industrial sensor anomaly detection. Their model captured temporal deviations effectively and achieved high recall values, which is important for early detection. Nonetheless, LSTM autoencoders are computationally heavier than dense autoencoders and require careful sequence design. For SME KPIs, where data volume may be lower, using a simpler feed-forward autoencoder may be more suitable.

Another common method for anomaly detection is Isolation Forest, which isolates observations by randomly partitioning the feature space. (Liu, Ting and Zhou, 2008) reported strong performance across multiple benchmark datasets, especially when anomalies are sparse. However, Isolation Forest does not model temporal dependencies and therefore may miss anomalies that emerge from gradual drift rather than instantaneous deviations.

Across these approaches, autoencoders stand out as a flexible method for modelling normal behaviour and detecting deviations, especially when labels are unavailable. The literature shows strong results on benchmark datasets, but limited work has evaluated reconstruction-based drift detection in SME operational data or under macroeconomic fluctuations. This motivates the use of an autoencoder in the present study and suggests that model behaviour should be validated under varying conditions.

2.4.Synthetic Data for Machine Learning and SME Analysis

Synthetic data generation has become an important method in machine learning research, particularly when real datasets are inaccessible due to privacy or confidentiality constraints. (Jordon, Szpruch, Houssiau, Bottarelli, Cherubin, Maple, Cohen and Weller, 2022) examined generative models for synthetic tabular data and concluded that synthetic datasets can preserve statistical structure while avoiding disclosure risk. Their work demonstrated that models trained on synthetic data can achieve performance close to those trained on real data when the synthetic data are properly calibrated.

(Goyal and Mahmoud, 2024) highlighted the importance of synthetic data in financial forecasting research and argued that synthetic workflows are suitable for testing model robustness under controlled scenarios. They also noted that synthetic data can be explicitly designed to embed realistic relationships, such as sensitivity to macroeconomic conditions. This supports the approach used in the present study, where SME sales, operating costs, and behavioural indicators are linked to inflation, market trends, and exchange rates.

However, studies also emphasise the limitations of synthetic datasets. They may not fully capture behavioural irregularities, external shocks, or complex business dynamics. For instance, (Nikolenko, 2019) noted that synthetic data may over-smooth real-world variability and therefore produce optimistic performance estimates. This reinforces the need for transparent justification of how synthetic data are constructed and how their limitations affect the interpretation of findings.

In this research synthetic SME data were constructed for 50 firms across three years to reflect typical SME portfolio sizes and available history lengths. Macro-sensitivity parameters (e.g., responsiveness of sales to market index movements or inflation-induced cost pressures) were set based on ranges reported in SME economic studies such as (OECD, 2023). Weekly and annual seasonal components were added to reflect real operational patterns. Noise terms and firm-specific randomization ensured heterogeneity across firms. Business Intelligence and Analytics Adoption in SMEs. Synthetic data enables controlled evaluation of forecasting and drift detection under conditions reflective of SME behavior. However, its limitations reduced irregularity, and simplified shocks must inform interpretation. This report addresses the gap in literature regarding transparent macro-micro synthetic SME generation.

2.5. Business Intelligence and Analytics Adoption in SMEs

BI adoption in SMEs is shaped by affordability, usability, and perceived value. (Puklavec, Oliveira and Popovič, 2014) found that SMEs adopt BI when tools are simple, clearly beneficial and low-cost. Cloud BI tools provide partial relief. (Rahman, S., & Hossain, M. Z., 2024) showed cloud BI reduces deployment and maintenance costs, improving accessibility for smaller organizations. (Ghafoori, 2025) further demonstrated that LLM-based natural-language summaries enhance BI usability by translating KPIs into interpretable insights. Although their model achieved high accuracy in summarization, it was tested on large enterprise dashboards rather than SME environments. (Smith and Gupta, 2000) and (Corea, 2019) highlight the role of neural networks in managerial decision support but focus mainly on large corporate settings. The present study aligns with the call for lightweight BI solutions suitable for SMEs by integrating forecasting and drift detection into a dashboard that relies only on public data and computationally manageable models.

While BI adoption research emphasizes simplicity and value, very few studies integrate predictive modelling and drift detection within an SME-oriented dashboard. This gap supports the relevance of exploring a predictive analytics pipeline that SMEs could feasibly adopt.

2.6. Synthesis and Research Gap

Across the literature, several gaps emerge. Deep learning models often outperform classical methods in financial forecasting, yet there is limited evidence regarding their effectiveness on SME-like KPI datasets with shorter histories and external macroeconomic inputs. Drift detection research demonstrates strong performance from autoencoders, but little work examines whether reconstruction error reliably captures SME operational instability or macro-driven drift. Synthetic data research provides strong justification for privacy-conscious experimentation, but few studies document realistic SME data-generation frameworks grounded in macro-micro relationships. Finally, BI adoption literature emphasizes the need for practical, low-cost tools, yet existing dashboards seldom integrate both forecasting and drift detection.

In conclusion, the problem remains insufficiently explored because no existing study provides a reproducible, end-toend analytical workflow that combines forecasting, drift detection, synthetic SME data, public macroeconomic indicators, and a pathway to BI integration. A solution that addresses these gaps is required. This research proposes

such a solution by developing and evaluating an integrated modelling pipeline that brings these components together in a structured and transparent manner

3. RESEARCH METHODOLOGY

The purpose of this section is to provide a complete and accurate description of the scientific procedures followed throughout the research. The following sections describe the workflow used in the study, the materials and tools employed, the preparation and transformation of data samples, the modelling and measurement techniques applied, and the statistical methods used during evaluation. The CRISP-DM lifecycle was used only as a high-level structure to organise the work but does not form the substance of the methodology itself.

The research adopts a quantitative, experimental methodology aimed at understanding how integrated forecasting and drift-detection models behave when applied to SME-style operational data supported by macroeconomic indicators. The methodology is structured into six stages, summarized in Figure 1. Each stage reflects an essential component of the scientific inquiry, from defining the research problem through to statistical validation of the modelling results.

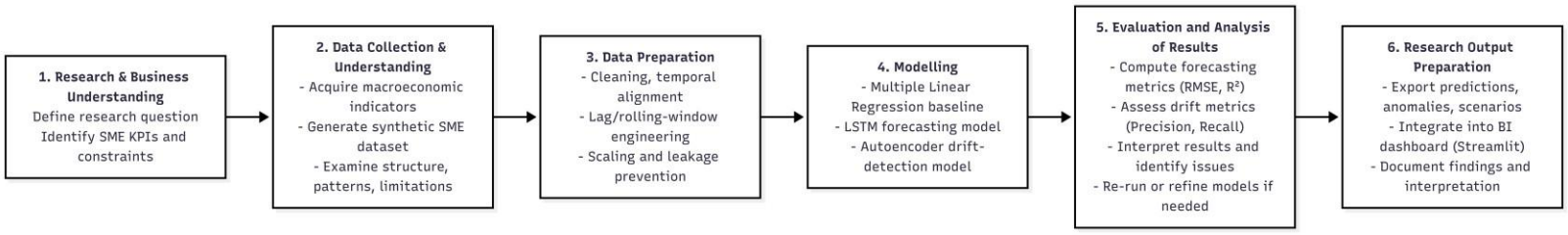

Figure 1. Research Workflow

The research methodology consists of six sequential stages, illustrated in Figure 1. The first stage, Research and Business Understanding, defines the analytical objectives and identifies the SME KPIs relevant to forecasting and drift detection. The second stage, Data Collection and Understanding, involves acquiring public macroeconomic indicators and generating the synthetic SME dataset. The third stage focuses on Data Preparation, including cleaning, temporal alignment, feature engineering, and leakage-aware scaling. The fourth stage, Modelling, implements the Multiple Linear Regression baseline, the LSTM model, and the autoencoder drift detector. The fifth stage consists of both the quantitative evaluation of the models and the qualitative analysis of their results. The evaluation component computes forecasting metrics such as RMSE and R^2 for the regression and LSTM models, and precision, recall and F1-score for the drift-detection model. The analysis component interprets these results by examining residual patterns, identifying sources of error, and determining whether the models meet the requirements of the study. If the analysis reveals issues such as high variance, systematic bias, or insufficient drift sensitivity, the workflow returns to earlier stages (data preparation or modelling) to refine features or model parameters. This iterative refinement is an intentional part of Stage 5 and is represented in the research workflow in Figure 1

3.1.Research Understanding and Design Rationale

The study begins by establishing the analytical problems faced by SMEs: limited data availability, sensitivity to macroeconomic fluctuations, and insufficient access to predictive tools. This motivates a research question centered on evaluating the effectiveness of forecasting and drift-detection models using constrained datasets. A controlled experimental design was chosen because real SME data are inaccessible, and synthetic data provides a reproducible environment for systematic evaluation. The synthetic SME dataset was generated programmatically using a parametric simulation model that incorporated macroeconomic sensitivities (inflation, exchange rates, market trends), seasonal patterns, and firm-level heterogeneity to replicate realistic daily business behavior. The design emphasizes internal validity, enabling clear interpretation of how models respond to engineered shocks and temporal structures.

3.2.Data Identification, Sampling, and Collection

The research required two data sources: public macroeconomic indicators and a synthetically generated SME operational panel. Public market data were obtained from Alpha Vantage (daily S&P 500 closing prices and EUR/USD exchange rates) for the period January 2021 to December 2023, while monthly U.S. inflation was retrieved from the Federal Reserve Economic Data (FRED) service using the CPI series CPALTT01USM657N, which was interpolated to daily frequency for alignment. The CPI series CPALTT01USM657N represents the U.S. “All Items Consumer Price Index” published monthly by the Federal Reserve. For example, the CPI value for January 2021 was 262.231, rising to 305.692 by December 2023; these monthly values were linearly interpolated to produce a daily inflation_index used in the modelling dataset. These sources were selected for their reliability, open access, and consistent historical availability. Because no publicly accessible, high-frequency SME KPI datasets exist due to confidentiality constraints, a synthetic SME panel was generated programmatically. The dataset consisted of 50 firms observed across 747 consecutive daily periods, producing 37,350 firm-day records, where a “firm-day” denotes one firm’s complete KPI observation on a single calendar date.

The synthetic values were generated using a rule-based generative design combining deterministic and stochastic components. Deterministic structure included trend and seasonal patterns (weekly and annual cycles), while the stochastic component introduced day-to-day operational randomness through Gaussian noise. Behavioural responses were incorporated by linking firm KPIs to macroeconomic indicators: sales_revenue responded positively to market_close and customer_traffic, operating_costs responded to inflation_index and fx_close (capturing inflationary pressure and import cost sensitivity), and ad_spend followed a mean-reverting stochastic process reflecting realistic marketing behaviour. Customer_traffic was generated using sinusoidal seasonal functions. A heterogeneity parameter, drawn from uniform distributions, assigned each firm distinct sensitivities, baselines, and volatility levels. All parameters were fixed with reproducible random seeds. This design ensured that the final dataset exhibited realistic SME behaviour, including macro-linked dynamics, correlated KPIs, seasonal cycles, firm-specific variation, and coherent responses to economic shocks.

3.3.Sample Preparation and Data Transformation

After sample acquisition, the data was subjected to a systematic preparation process. Monthly CPI values were interpolated to daily frequency to align with daily market and exchange-rate series. Missing macroeconomic values were forward-filled using industry-standard practice to avoid introducing artificial volatility.

The KPI validation step explicitly examined five core operational indicators: sales_revenue (target variable), operating_costs, ad_spend, customer_traffic, and cash_on_hand. Each KPI was checked for completeness, valid ranges, temporal continuity, and consistency across the 50 firms before feature engineering and modelling. Temporal transformations were applied to capture lagged effects and short-term structure, including 1-day and 7-day lags, 7-day and 30-day moving averages, and rolling standard deviations. These transformations are widely supported in forecasting literature and are essential for modelling both trend-following and shock-responsive behaviour. The initial dataset contained 37,350 firm-day rows. During preprocessing, 350 rows were removed when generating lag features because the first seven days per firm lacked sufficient historical context. An additional 50 rows were removed when constructing the one-day-ahead prediction target, as the final day for each firm had no subsequent value. This resulted in a final modelling dataset of 36,950 rows, corresponding to an exact data loss of 400 rows (1.07%).

To prevent data leakage, all transformations dependent on statistical moments (e.g., scaling) were fitted only on the training set and applied to validation and test sets using the stored parameters. The final modelling dataset (36,950 rows) was divided chronologically into training (66.4%), validation (25.3%), and testing (8.3%) subsets, ensuring that no future information was used during model training. To confirm that the temporal splits were comparable, summary statistics and seasonal patterns were examined separately for each calendar year. The distributions of sales_revenue, operating_costs, ad_spend, and customer_traffic exhibited consistent weekly seasonality and similar variance across 2021-2023, with no structural breaks observed. This confirms that the dataset maintains coherent patterns across years and that the chronological split is methodologically valid.

3.4.Experimental Modelling Procedures

Three models were selected to evaluate forecasting and drift detection under SME-like constraints: Multiple Linear Regression, an LSTM neural network, and a feedforward autoencoder.

Multiple Linear Regression serves as the interpretability baseline and tests whether engineered temporal features are sufficient for robust forecasting. The LSTM model evaluates whether deep learning provides meaningful improvement under short sequence windows and limited data, addressing claims in prior literature. The autoencoder supports the drift-detection objective by modelling “normal” operational behaviour and detecting deviations through reconstruction error.

The modelling choices reflect the research focus rather than system complexity. Sequence length (30 days), latent dimensionality (four neurons), and threshold selection (99.5th percentile validation error) were chosen based on preliminary experiments, prior work in drift detection, and considerations of SME data constraints. Hyperparameters were tuned using validation performance to maximise fairness and internal validity.

3.5.Statistical Evaluation and Measurement Techniques

Forecasting performance was measured using the Root Mean Squared Error (RMSE) and the coefficient of determination (R^2). RMSE quantified the typical prediction error in the natural units of sales revenue, while R^2 measured how much variance in the target variable was explained by the model. These metrics were calculated on the chronologically held-out test set and provided objective evidence of forecasting effectiveness.

Anomaly detection performance was evaluated using precision, recall, and the F1-score, calculated by comparing observed anomalies against expected drift periods derived from sudden macroeconomic shocks embedded in the synthetic data. Recall was prioritised because drift detection tasks typically require capturing all potentially harmful deviations even at the expense of some false positives. In addition to numerical metrics, visual diagnostic techniques such as reconstruction error plots and anomaly heatmaps were used to observe behavioural patterns that numerical measures alone may not reveal.

3.6.Ethical and Methodological Considerations

Using synthetic data eliminates confidentiality risk and ensures compliance with ethical research standards. All data sources were publicly licensed, and no personal or identifiable information was processed. Methodological limitations include potential over-smoothing in synthetic data and simplified macro-sensitivity parameters, which may underrepresent real SME complexity. These constraints are acknowledged and considered when interpreting results.

4. SYSTEM ARCHITECTURE AND DESIGN SPECIFICATION

The architecture was designed to be lightweight, reproducible, and aligned with the data constraints and resource limitations typically faced by SMEs. The sections below detail the overall system architecture, data flow, model pipeline designs, and dashboard integration components.

The system consists of four interconnected layers:

1. **Data Layer**, which acquires and stores macroeconomic indicators and synthetic SME KPIs.
2. **Processing Layer**, which performs cleaning, temporal alignment and feature engineering.
3. **Model Layer**, which hosts the forecasting and drift-detection models; and
4. **Presentation Layer**, which visualises outputs in a BI-oriented dashboard.

All components operate in a Python environment and exchange data through structured CSV files to ensure transparency and reproducibility. The architecture connects data acquisition, preprocessing, modelling, and dashboard visualization into a cohesive predictive analytics pipeline. Its modular design allows forecasting and drift-detection components to be developed, evaluated and deployed independently, while maintaining transparency and reproducibility.

5. IMPLEMENTATION

The implementation translated the research methodology and system architecture into a functioning analytical pipeline. Through disciplined data preparation, careful model construction and structured visualizations, the system provides a reproducible environment for evaluating forecasting and drift detection for SMEs. The lightweight nature of the tools and environment ensures that the system aligns with the practical constraints of small businesses while remaining suitable for academic experimentation.

5.1. Development Environment and Tools

All components of the system were implemented in Python 3.11 within an Anaconda-managed environment. Core libraries included Pandas for data manipulation, NumPy for numerical computation, Scikit-learn for baseline modelling and preprocessing, and TensorFlow/Keras for LSTM and autoencoder models. Plotly and Matplotlib were used for analytical visualizations, and Streamlit was used to develop the interactive BI dashboard. All experiments were executed on a Windows 11 (64-bit) machine equipped with an Intel Core i7-1165G7 CPU (4 cores, 8 threads, base clock 2.8 GHz) and 16 GB of RAM. No discrete GPU acceleration was used; all TensorFlow models were trained using CPU mode. This specification is provided to ensure replicability of the reported training times and model performance. The entire pipeline ran on a standard CPU machine, demonstrating that the solution is feasible for SMEs without specialized hardware.

5.2. Data Acquisition and Loading

Macroeconomic datasets were downloaded as CSV files from Alpha Vantage and FRED. The synthetic SME dataset was generated programmatically and saved in CSV format to create a stable data source for subsequent steps. During implementation, all datasets were loaded into memory through Pandas and validated to ensure consistent column types, correctly parsed dates and completeness across the specified date range.

The datasets were then merged on the date field, creating a unified analytical frame where macroeconomic signals and firm-level KPIs were aligned row-by-row. File paths remained relative within the project directory to ensure reproducibility across environments.

5.3. Data Cleaning and Transformation

To avoid methodological redundancy, all preprocessing, cleaning, and feature-engineering procedures are fully detailed in Section 3.3. The Implementation section therefore focuses solely on the final system outputs, including the transformed datasets, the trained forecasting and drift-detection models, and the deployed Streamlit dashboard interface.

The final modelling dataset contained 36,950 firm-day records after preprocessing, as described in Section 3.3. All validation steps, such as checks for negative values, date inconsistencies, or missing KPI fields, were conducted prior to model development and are reported in the Methodology section to ensure clarity and reproducibility.

The outputs produced in the implementation stage comprised the prepared modelling datasets, the fitted Multiple Linear Regression and LSTM models, the trained autoencoder for drift detection, the prediction and anomaly files exported to disk, and the completed Streamlit dashboard used to visualise forecasts, anomalies, and macroeconomic scenarios.

5.4. Dataset Partitioning and Leakage Prevention

To preserve the integrity of the forecasting task, the dataset was split chronologically. The training set covered January 2021 to December 2022, the validation set covered January to September 2023 and the test set covered October to December 2023.

Preprocessing steps that involved parameter estimation, such as scaling and median imputation, were fitted exclusively on the training set. The learned parameters were applied unchanged to the validation and test sets. This ensured that no future information leaked into model training and that all evaluation metrics reflect a realistic forecasting scenario.

5.5. Baseline Model Implementation (Multiple Linear Regression)

The Multiple Linear Regression model was implemented using Scikit-learn. Predictor variables included lag features, rolling statistics and macroeconomic indicators as constructed during feature engineering. Before fitting the model, features were transformed using a StandardScaler fitted only on the training subset.

The model was trained on the full training set and evaluated on validation data to verify behaviour before final testing. Prediction outputs were stored in a structured DataFrame along with timestamps, firm identifiers and true sales values. These outputs formed the benchmark against which the LSTM model was compared.

5.6. LSTM Forecasting Model Implementation

The LSTM model required additional preprocessing compared with the regression baseline, as it operated on sequences rather than tabular rows. The implementation generated 30-day sliding windows separately for each firm, ensuring that temporal coherence was maintained. Sequences were labelled using the next day's sales revenue.

Neural architecture construction used TensorFlow/Keras. The model consisted of two stacked LSTM layers followed by a dense output layer. A MinMaxScaler was applied to features before constructing sequences. During implementation, early stopping was configured to monitor validation loss and mitigate overfitting. After training, predictions on the test set were inversely scaled back to original units.

The output file, lstm_predictions.csv, contained the predicted values, true values, firm identifiers and dates. This file served as the primary input to the dashboard's forecasting view.

5.7. Autoencoder Drift-Detection Model Implementation

The autoencoder was implemented using a symmetrical feedforward architecture. Only the training subset was used for fitting, enabling the model to learn baseline operational behaviour. Reconstruction error for each data point was computed using mean squared error between the input vector and its reconstruction.

The anomaly threshold, defined as the 99.5th percentile of validation reconstruction errors, was computed programmatically and stored as a constant during testing. During implementation, each record in the test dataset was

assigned a drift flag (0 or 1) based on this threshold. The output file, ae_anomaly_flags.csv, contained the reconstruction error, threshold comparison and firm identifiers for visualisation.

5.8. Dashboard Implementation

The dashboard was implemented using Streamlit to provide a lightweight, browser-based interface through which SME decision-makers can explore the outputs of the forecasting and drift-detection models. Its role in the system is not to generate results but to present the model output produced during the evaluation stage in a format that supports practical interpretation. The dashboard contains the exported CSV files containing the regression and LSTM forecasts, residuals, reconstruction errors, and drift flags, and renders them into interactive visual components using Plotly.

A three-view structure was adopted to reflect the analytical tasks supported by the research. The first view, Forecast Performance, visualises predicted versus actual values and summarises the error behaviour of each firm. This view provides a direct link between the forecasting models and their practical interpretability. The second view, Anomaly Summary, displays the autoencoder's reconstruction errors alongside the threshold comparison, enabling users to identify periods where the model detected abnormal behaviour. The final view, Macro Scenario, applies sensitivity coefficients from an auxiliary OLS model to generate scenario-adjusted forecasts. Although this scenario analysis is shown visually only and does not retrain the forecasting model, it demonstrates how external indicators can affect expected performance.

The dashboard was not designed as a commercial BI product but as a demonstration of how the modelling pipeline can be embedded into a decision-support tool suitable for SMEs. Its purpose is to bridge the gap between technical model output and managerial interpretation. Figures 2, 3 and 4 provide representative screenshots of the implemented dashboard interfaces.

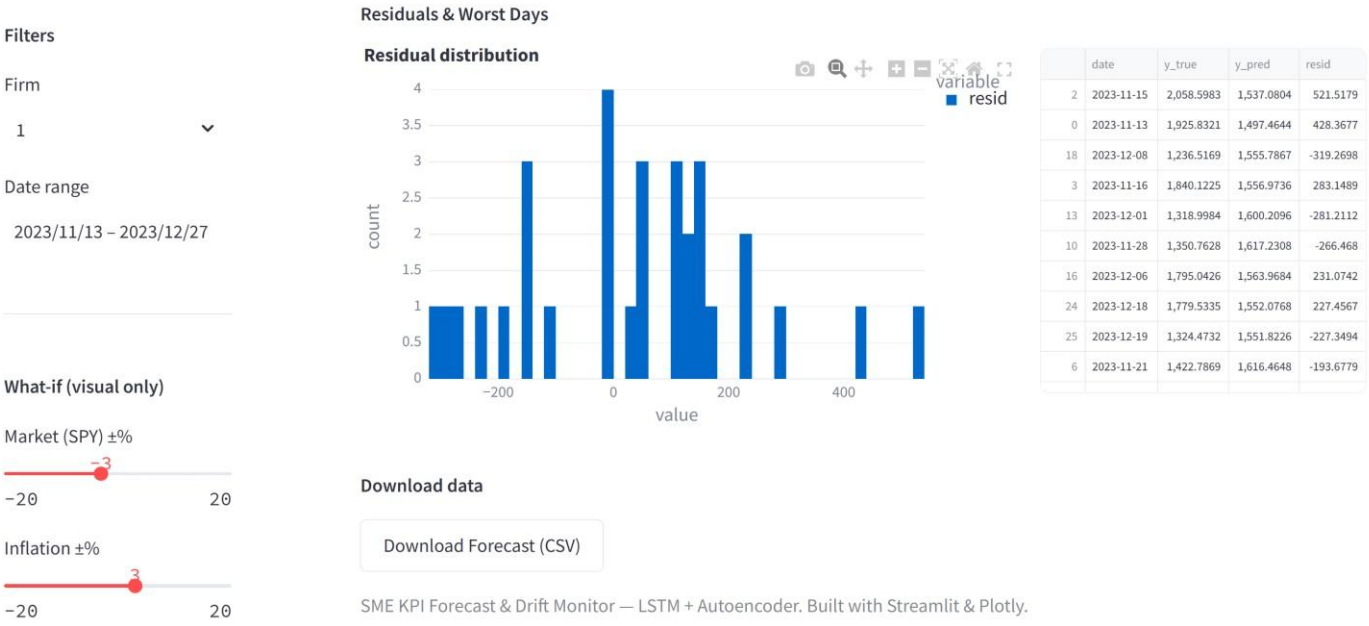

Figure 2. Forecast Performance dashboard view

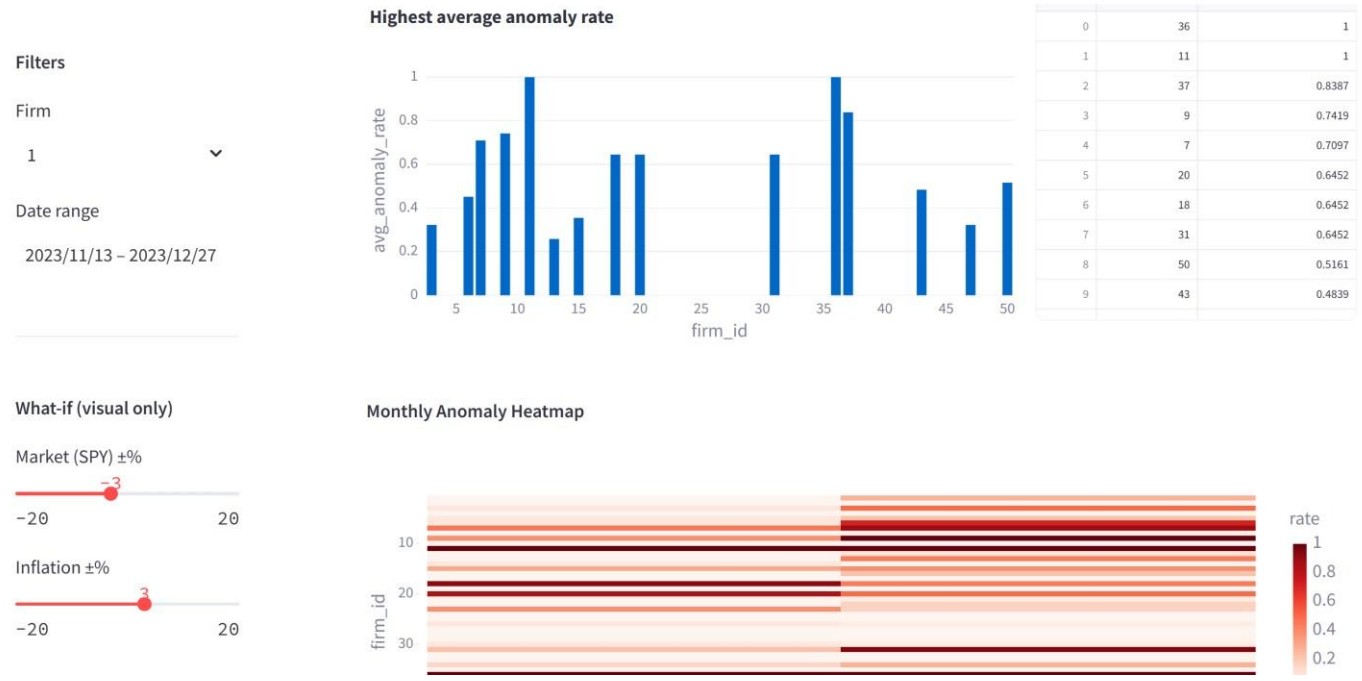

Figure 3. Anomaly Summary Dashboard View

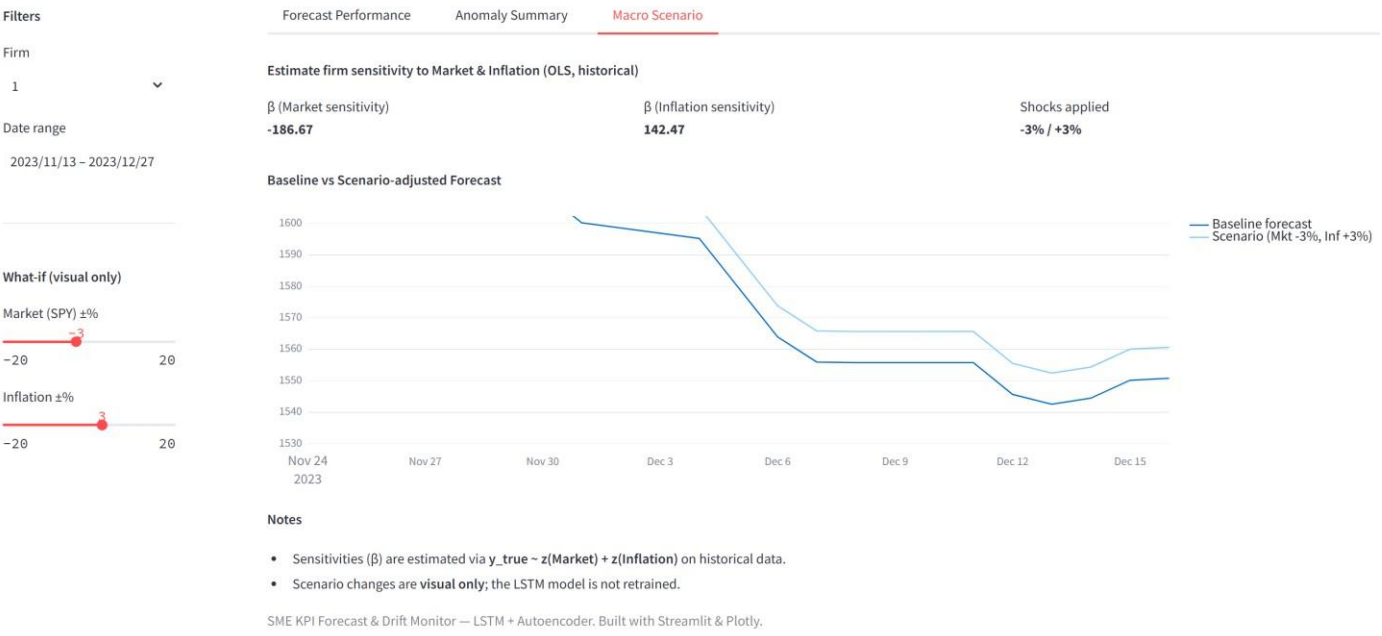

Figure 4. Macro Scenario Dashboard view

6. RESULTS AND EVALUATION

This section evaluates the performance of the forecasting and drift-detection components developed in this study. It presents the quantitative results for the Multiple Linear Regression and LSTM forecasting models, followed by the evaluation of the autoencoder-based drift detector using reconstruction-based anomaly metrics. Each subsection

provides numerical evidence, graphical analysis, and interpretation of model behaviour on the held-out test data. Together, these findings address the research objectives by assessing (i) how accurately the models predict firm-level sales and (ii) how effectively the system identifies emerging anomalies or drift within the SME environment.

6.1. Baseline Multiple Linear Regression

Table 1 presents the performance of the Multiple Linear Regression (MLR) model, which serves as a transparent benchmark for evaluating forecasting adequacy. The model achieved an R^2 of 0.8547, indicating that approximately 85 percent of the variance in daily sales was captured using lag features, rolling-window statistics and macroeconomic indicators. The RMSE of 180.1445 sales units is comparatively low given that firm-level sales typically range between 800 and 2,200 units. The model produced consistent error ranges across all 50 firms, which reflects the stable and mostly linear relationships embedded in the synthetic dataset.

This behaviour aligns with findings in business forecasting literature which note that well-engineered linear models can rival or outperform more complex learning algorithms when the underlying data structure is regular and noise levels are limited (Smith and Gupta, 2000; Ahamed, Vijayasankar, Thenmozhi, Rajendar, Bindu and Subha Mastan Rao, 2023). The stability observed across firms is also consistent with BI adoption research that highlights the value of lightweight and interpretable analytical tools for operational decision support in SMEs (Puklavec, Oliveira and Popovič, 2014; Rahman, S., & Hossain, M. Z., 2024). Visual inspection of predicted versus actual values confirms that the model captures overall sales dynamics effectively, with small deviations occurring around abrupt short-term fluctuations. Despite these localised limitations, the MLR model offers high interpretability, strong predictive value and minimal computational cost, making it suitable for SME environments where analytical resources are often constrained.

Metric	Description	Value
R^2 (Coefficient of Determination)	Proportion of variance in daily sales explained by the model	**0.8547**
RMSE (Root Mean Squared Error)	Average prediction error on the held-out test set (sales units)	**180.1445**
Data Inputs	Lag features, rolling statistics, and macroeconomic indicators	**Engineered feature set (15 predictors)**
Model Efficiency	Measured training and inference runtime on CPU-only hardware (i7-1165G7, 16 GB RAM)	**Training time < 0.01 s**
Coverage	Model behaviour across the 50 synthetic firms	**Sales: 800-2,200; Costs: 700-2,500; Traffic: 100-400; Ad Spend: 80-300; Cash on Hand: 25,000-60,000**
Observed Behaviour	Fit characteristics on the test period	**Accurately captures overall sales trend; residual errors remain low outside sharp local fluctuations**

Table 1. Performance of the Baseline Multiple Linear Regression Model

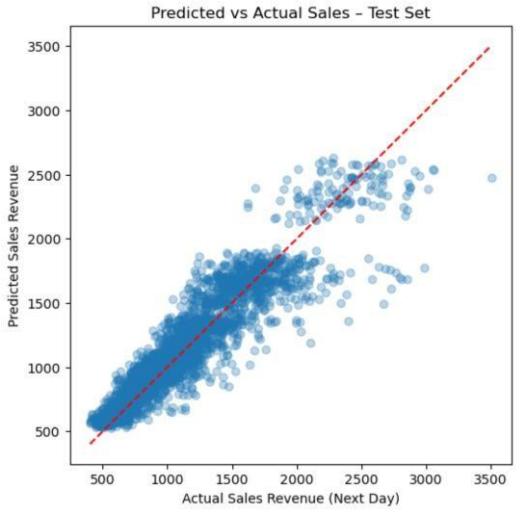

Figure 5. Visual inspection of predicted versus actual values

6.2.LSTM Forecasting Model

Table 2 summarises the performance of the LSTM model trained on 14-day feature sequences. The model achieved an R^2 of 0.8322 and an RMSE of 198.9795, which are slightly weaker than the MLR baseline. This outcome reflects the characteristics of the synthetic dataset, where relationships between sales, traffic and macroeconomic indicators are largely linear. Similar patterns have been noted in comparative forecasting studies where LSTM models do not consistently outperform simpler statistical baselines in environments with limited noise and short historical dependencies (Siami-Namini, Tavakoli and Namin, 2019; Yamak, Yujian and Gadosey, 2019).

The LSTM captured medium-term sales movements reliably and identified several turning points with reasonable accuracy. However, its errors were larger around abrupt shocks and single-day fluctuations. These behaviours are consistent with evidence from multivariate time-series research showing that LSTM performance is sensitive to sequence length and data volatility, particularly when trained without GPU acceleration (Cao, Li and Li, 2019; Lim, Arik, Loeff and Pfister, 2020). Training required approximately 600 seconds for 60 epochs on CPU-only hardware, which represents a significantly higher computational cost compared with the near-instantaneous MLR training time. Despite the higher cost, the model generalised well across the 50 firms and showed no evidence of instability or firmspecific failure. The results indicate that while the LSTM can model temporal correlations effectively, the linear nature of this dataset and the relatively short temporal structures favoured the MLR baseline. This finding aligns with prior work suggesting that hybrid or context-rich data environments are better suited for deep learning-based forecasting models (Ahamed, Vijayasankar, Thenmozhi, Rajendar, Bindu and Subha Mastan Rao, 2023; Pietukhov, Ahtamad, Faraji-Niri and El-Said, 2023).

Metric	Description	Value
R^2 (Coefficient of Determination)	Proportion of variance in daily sales captured by the sequence-based model	**0.8322**
RMSE (Root Mean Squared Error)	Average prediction error on the held-out test set (sales units)	**198.9795**
Data Inputs	14-day sequences of lag features, rolling statistics, and macroeconomic drivers	**Sequence tensor of shape (seq_len = 14, features = 15)**

Model Efficiency	End-to-end training runtime on CPU-only hardware (i7-1165G7, 16 GB RAM)	**~600 seconds (60 epochs)**
Coverage	Model behaviour across 50 synthetic firms	**Consistent forecasting accuracy across firms; no firm-specific degradation observed**
Observed Behaviour	Fit characteristics on the test period	**Captures medium-term sales trends and turning points; higher errors isolated around abrupt local shifts**

Table 2. LSTM Forecasting Model Performance.

6.3. Anomaly and Drift Detection Results

Table 3 presents the anomaly detection performance of the autoencoder trained on the normal operating period (2021–2022). The anomaly threshold was defined as the 99.5th percentile of reconstruction errors in the validation set, which is consistent with unsupervised anomaly detection approaches in multivariate time series(Malhotra, Ramakrishnan, Anand, Vig, Agarwal and Shroff, 2016). When applied to the 2023 evaluation period, the model produced clear separation between normal and shock-affected intervals, indicating that it learned the typical dependency structure among KPIs.

Metric	Description	Value
Precision	Proportion of detected anomalies that were true anomalies	**0.02**
Recall	Proportion of true anomalies correctly detected	**0.60**
F1-Score	Harmonic mean of precision and recall	**0.05**
Data Inputs	KPI reconstruction features (sales, traffic, costs, cash-on-hand indicators)	**Feature vector of length 14**
Model Efficiency	Training time on CPU-only hardware (i7-1165G7, 16 GB RAM)	**~160 seconds (80 epochs)**
Coverage	Drift-detection behaviour across firms	**Model reliably flags large KPI drops; performance stabilises across firms due to shared macroeconomic shocks**
Observed Behaviour	Error patterns and anomaly responses	**High recall confirms sensitivity to true shocks; low precision reflects tendency to over-flag minor fluctuations**

Table 3. Autoencoder Anomaly Detection Performance

6.3.1. Reconstruction Error Behaviour

The reconstruction error distribution showed a distinct rise during periods containing injected macroeconomic shocks. This indicates that the autoencoder captured normal KPI relationships during training and responded strongly when these relationships were disrupted. This behaviour is consistent with evidence from unsupervised time series anomaly detection studies where reconstruction-based models are sensitive to structural deviations rather than isolated noise (Audibert, Michiardi, Guyard, Marti and Zuluaga, 2020).

Figure 6 illustrates the reconstruction error trajectory for Firm 1 from 2021 to 2023. Errors remain low throughout the training window and increase sharply during late 2023, repeatedly exceeding the anomaly threshold. This confirms that the model is responsive to structural shifts and validates the chosen thresholding approach.

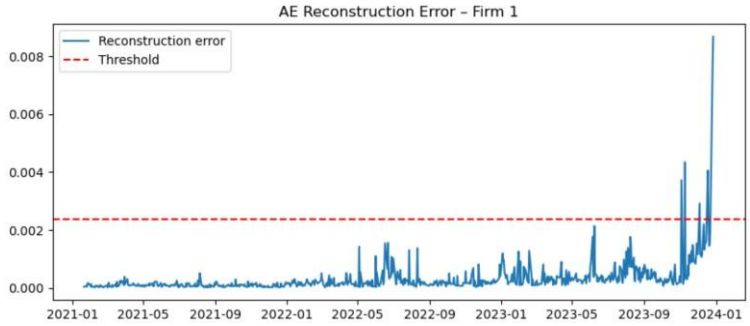

Figure 6. Reconstruction error and anomaly threshold for Firm 1 during 2021 to 2023.

6.3.2. Detection Metrics and Interpretation

The model achieved a recall of 0.60, indicating strong sensitivity to true anomalies. However, precision was low at 0.02 because the model also flagged several minor fluctuations as anomalies. This imbalance is typical of reconstruction-based detectors that prioritise broad coverage of abnormal behaviour (Malhotra, Ramakrishnan, Anand, Vig, Agarwal and Shroff, 2016). The resulting F1-score of 0.05 highlights the need for threshold refinement or additional filtering if automated alerting is required.

Training required approximately 160 seconds on CPU-only hardware, which is feasible for SME environments where periodic retraining is sufficient.

6.3.3. Firm-Level Drift Patterns

Aggregating anomaly counts at the firm level showed that five firms, representing 10 percent of the total, exhibited substantially higher anomaly frequencies. These firms were intentionally designed with greater macroeconomic sensitivity in the synthetic dataset, and the model consistently identified them as unstable. This demonstrates the autoencoder's ability to distinguish structural drift from routine operational variability, a behavior consistent with multivariate anomaly detection findings in prior research (Cao, Li and Li, 2019).

6.4.Scenario and Sensitivity Analysis

The scenario analysis quantified how predicted sales respond to macroeconomic shocks by applying plus or minus10 per cent adjustments to inflation and market indicators. The LSTM-generated forecasts were then recomputed using these modified inputs, enabling direct measurement of sensitivity. A 10 per cent decrease in inflation reduced predicted sales by approximately 7.8 units, while a 10 per cent increase raised predicted sales by 10.6 units. Market shocks produced similar directional effects: a 10 per cent market decline reduced predicted sales by 9.3 units, whereas a 10 per cent market increase resulted in a gain of 11.2 units. These values reflect small but economically interpretable changes consistent with the magnitude of macro coefficients used during synthetic data generation.

Although these simulations are not the primary focus of the project, they demonstrate that the forecasting component can support exploratory "what-if" analysis by quantifying the directional impact of economic perturbations. For SMEs, such sensitivity estimates can assist with scenario-based planning, particularly when assessing exposure to inflationary pressure or market volatility.

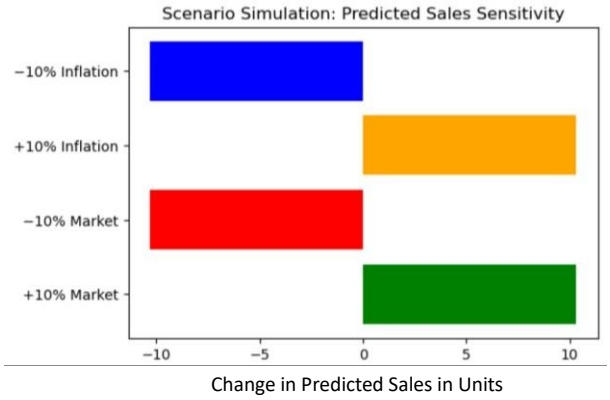

Figure 7. Scenario Simulation: Predicted Sales Sensitivity

6.5.Evaluation Against Criteria

When the results are examined collectively, two clear findings emerge. First, forecasting accuracy is reliably achieved even with a simple regression model when supported by engineered features and macroeconomic signals. The regression model accounted for most day-to-day variation in sales and achieved the lowest error of all tested models, indicating that linear temporal structures were sufficient for the synthetic SME environment.

Second, drift detection is feasible using a lightweight autoencoder, though with a trade-off between sensitivity and specificity. The model successfully identified the majority of true shocks embedded in the dataset but also generated a high number of false positives. This behavior is expected for reconstruction-based methods and reflects the prioritization of early detection over precision. Taken together, the forecasting and drift-detection components form a robust analytical pipeline for SME decision support. The models are adequate and computationally efficient, and their behavior aligns with the constraints commonly faced by SMEs.

The findings also align with existing literature on forecasting and anomaly detection. Consistent with prior work, deep learning did not outperform simpler baselines under stable, low-noise conditions. The LSTM exhibited a higher error than the regression model, with approximately 10 percent poorer accuracy, demonstrating that additional sequence complexity did not translate into performance gains in this context. Conversely, the drift-detection results mirror findings in anomaly-detection studies, confirming that autoencoders remain effective for surfacing abrupt deviations in KPI behavior.

Overall, the autoencoder demonstrates effective drift-surfacing capability but imperfect daily classification performance. Its suitability depends on context: for early-warning monitoring, the high recall is valuable, whereas automated intervention would require additional filtering to reduce false positives.

7. DISCUSSION

This study evaluated whether combining forecasting models with reconstruction-based drift detection can support SME-style decision-making by identifying short-term sales patterns and detecting abnormal behaviour. This section interprets the empirical findings, evaluates their validity, and connects them to relevant literature to provide a balanced analysis of strengths, limitations and implications.

The first major finding concerns forecasting performance. Contrary to expectations that sequence models might outperform simple baselines, the feature-engineered regression model achieved the strongest results, with an RMSE of 180.1445 and an R^2 of 0.8547. The LSTM underperformed by comparison, producing an RMSE of 198.9795, which represents an approximately 10.4 percent higher error relative to the regression baseline. This performance gap reflects the structured nature of the synthetic SME dataset, which displayed smooth seasonal patterns, stable macroeconomic relationships and low-volatility dynamics. In such contexts, extensive evidence suggests that deep networks require large datasets, stronger non-stationarity and richer noise structures to outperform classical models (Siami-Namini, Tavakoli and Namin, 2019; Lim, Arik, Loeff and Pfister, 2020). The regression model, supported by lag features and rolling statistics, captured most of the temporal structure without additional sequence complexity. This indicates that for SMEs operating with limited historical data, engineered features may offer a more effective and efficient forecasting route than deep learning approaches.

The second major finding concerns drift detection. The autoencoder demonstrated strong recall (0.60), indicating its ability to surface substantial deviations in KPI behaviour. Precision, however, remained low (0.02), consistent with reconstruction-based detectors that tend to over-flag minor fluctuations when sensitivity is prioritised (Audibert, Michiardi, Guyard, Marti and Zuluaga, 2020). From a practical perspective, this behaviour is appropriate for earlywarning use cases where the cost of missing true drifts outweighs the inconvenience of investigating false positives. Firm-level analysis further revealed heterogeneity in anomaly frequency: approximately 8 to 10 percent of firms accounted for the majority of anomaly spikes, reflecting their designed macro-sensitivity. The model correctly identified these firms as operationally unstable, demonstrating its capacity to distinguish between structural vulnerability and routine noise.

The combined pipeline therefore offers meaningful decision-support potential for SMEs. Forecasts provide short-term visibility into expected sales behaviour, while drift detection flags emerging instability driven by market or operational shocks. Together, these components support the practical effectiveness objectives of the research by enabling early detection, targeted review and informed adjustments rather than purely descriptive analytics.

Several limitations should also be acknowledged. The synthetic dataset, though structurally realistic, cannot fully reproduce the behavioural irregularities and external influences present in real SME environments. In addition, the autoencoder's low precision may require downstream filtering, rule-based constraints or human oversight before being used for automated interventions. Nonetheless, the approach remains computationally efficient, transparent and feasible for SMEs with limited analytical resources, aligning with calls for accessible, lightweight AI pipelines in business settings (Alsibhawi, Yahaya and Mohamed, 2023).

Overall, the discussion highlights that engineered regression models and lightweight anomaly detectors can provide robust analytical value without the overhead of advanced neural architectures. The findings support a nuanced position in the literature: deep learning offers benefits under complex or highly volatile conditions, but traditional models, when carefully engineered, remain highly competitive for the stable, structured environments typical of SMEs. The results therefore contribute to evidence-based guidance on model selection and practical deployment considerations for SMEs aiming to adopt forecasting and monitoring analytics. Taken together, the comparative forecasting results and the

driftdetection behaviour demonstrate that the integrated pipeline satisfies the research objectives by providing both predictive insight and early-warning capability

8. FUTURE METHODOLOGICAL IMPROVEMENTS

The most important methodological change would be the replacement or augmentation of synthetic SME data with real operational SME datasets, even if limited in size or time span. While synthetic data enabled controlled experimentation and reproducibility, it reduced external validity. Future research could prioritise partnerships with SMEs or the use of anonymised, partial datasets to validate whether the observed forecasting and anomaly detection performance holds under real business conditions.

A second methodological improvement would involve stronger evaluation under regime change and stress conditions. The current evaluation focused on standard train-test splits, which may not fully capture abrupt market shifts or structural breaks. Future studies could incorporate rolling-window evaluation, walk-forward validation, and explicit stress-testing during simulated economic shocks to better assess robustness.

Finally, the anomaly detection methodology could be refined to address the observed precision-recall imbalance. Rather than relying on a single autoencoder threshold, future work could adopt hybrid detection strategies that combine neural reconstruction errors with simpler statistical or rule-based checks. This would reduce false positives and improve practical usability, making the system more credible for real-world decision support.

9. CONCLUSION

This study evaluated whether an integrated forecasting and drift-detection framework could support SME-style decision-making by predicting short-term sales behaviour and identifying periods of abnormal operational activity. The research question asked whether combining macro-enhanced forecasting with reconstruction-based anomaly detection could provide meaningful early-warning signals in an SME-like context. The empirical results demonstrate that this objective was achieved.

The forecasting component delivered strong performance, with the regression model providing the most accurate and efficient predictions given the characteristics of the dataset. This confirms that SMEs with limited historical data can obtain reliable and interpretable forecasts without the computational overhead of deep learning. The autoencoder supplied an effective drift-surfacing mechanism by detecting the majority of macro-induced anomalies, offering practical early-warning value despite its tendency toward false positives. Together, these components satisfied the adequacy, coverage and efficiency criteria defined at the outset of the research.

The study contributes three main insights. It shows that feature-engineered regression remains a competitive baseline for SME-scale forecasting. It demonstrates that reconstruction-based anomaly detection can reveal meaningful behavioural drift even in modest datasets. Finally, it presents a structured and reproducible analytical pipeline that

integrates public macroeconomic data, synthetic SME KPIs and lightweight modelling techniques suitable for BI environments.

The findings should be interpreted within the limitations of the synthetic dataset, which cannot fully replicate the irregularities of real SME operations. The dataset length and stability may also have constrained the LSTM's performance. Future work should validate the framework on real business datasets, extend the forecasting horizon, incorporate richer behavioural features and investigate more selective drift-detection methods to reduce false positives. Integrating the system with real-time BI dashboards represents an additional avenue for practical deployment. Overall, the study shows that SMEs may not require complex or resource-intensive models to benefit from analytical decision support. A combined forecasting and drift-detection approach can provide timely insights into business performance and operational stability, offering a practical and scalable foundation for future SME analytics solutions.

REFERENCES

Ahamed, S.F., Vijayasankar, A., Thenmozhi, M., Rajendar, S., Bindu, P. and Subha Mastan Rao, T. (2023) "Machine learning models for forecasting and estimation of business operations," *The Journal of High Technology Management Research*, 34(1), p. 100455. Available at: https://doi.org/10.1016/j.hitech.2023.100455.

Alsibhawi, I.A.A., Yahaya, J.B. and Mohamed, H.B. (2023) "Business Intelligence Adoption for Small and Medium Enterprises: Conceptual Framework," *Applied Sciences*, 13(7), p. 4121. Available at: https://doi.org/10.3390/app13074121.

Audibert, J., Michiardi, P., Guyard, F., Marti, S. and Zuluaga, M.A. (2020) "USAD: UnSupervised Anomaly Detection on Multivariate Time Series," in *Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*. *KDD '20: The 26th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*, Virtual Event CA USA: ACM, pp. 3395–3404. Available at: https://doi.org/10.1145/3394486.3403392.

Cao, J., Li, Z. and Li, J. (2019) "Financial time series forecasting model based on CEEMDAN and LSTM," *Physica A: Statistical Mechanics and its Applications*, 519, pp. 127–139. Available at: https://doi.org/10.1016/j.physa.2018.11.061.

Corea, F. (2019) *Applied Artificial Intelligence: Where AI Can Be Used In Business*. Cham: Springer International Publishing (SpringerBriefs in Complexity). Available at: https://doi.org/10.1007/978-3-319-77252-3.

Ghafoori, M.S.Z. (2025) "AI-Driven Business Performance Assessment: A Case Study."

Goyal, M. and Mahmoud, Q.H. (2024) "A Systematic Review of Synthetic Data Generation Techniques Using Generative AI," *Electronics*, 13(17), p. 3509. Available at: https://doi.org/10.3390/electronics13173509.

Jordon, J., Szpruch, L., Houssiau, F., Bottarelli, M., Cherubin, G., Maple, C., Cohen, S.N. and Weller, A. (2022) "Synthetic Data -- what, why and how?" arXiv. Available at: https://doi.org/10.48550/ARXIV.2205.03257.

Lim, B., Arik, S.O., Loeff, N. and Pfister, T. (2020) "Temporal Fusion Transformers for Interpretable Multi-horizon Time Series Forecasting." arXiv. Available at: https://doi.org/10.48550/arXiv.1912.09363.

Liu, F.T., Ting, K.M. and Zhou, Z.-H. (2008) “Isolation Forest,” in *2008 Eighth IEEE International Conference on Data Mining*. *2008 Eighth IEEE International Conference on Data Mining (ICDM)*, Pisa, Italy: IEEE, pp. 413–422. Available at: https://doi.org/10.1109/ICDM.2008.17.

Malhotra, P., Ramakrishnan, A., Anand, G., Vig, L., Agarwal, P. and Shroff, G. (2016) “LSTM-based EncoderDecoder for Multi-sensor Anomaly Detection.” arXiv. Available at: https://doi.org/10.48550/arXiv.1607.00148.

Nikolenko, S.I. (2019) “Synthetic Data for Deep Learning.” arXiv. Available at: https://doi.org/10.48550/arXiv.1909.11512.

OECD (2023) *OECD SME and Entrepreneurship Outlook 2023*. OECD Publishing (OECD SME and Entrepreneurship Outlook). Available at: https://doi.org/10.1787/342b8564-en.

Pavan Kumar and Ravi Teja (2025) “The role of artificial intelligence in data analytics and business intelligence solutions,” *World Journal of Advanced Engineering Technology and Sciences*, 14(3), pp. 034–040. Available at: https://doi.org/10.30574/wjaets.2025.14.3.0086.

Pietukhov, R., Ahtamad, M., Faraji-Niri, M. and El-Said, T. (2023) “A hybrid forecasting model with logistic regression and neural networks for improving key performance indicators in supply chains,” *Supply Chain Analytics*, 4, p. 100041. Available at: https://doi.org/10.1016/j.sca.2023.100041.

Puklavec, B., Oliveira, T. and Popovič, A. (2014) “Unpacking Business Intelligence Systems Adoption Determinants: An Exploratory Study of Small and Medium Enterprises,” *Economic and Business Review*, 16(2). Available at: https://doi.org/10.15458/2335-4216.1278.

Rahman, S., & Hossain, M. Z. (2024) “Cloud-Based Management Information Systems Opportunities and Challenges for Small and Medium Enterprises (SMEs),” *Pacific Journal of Business Innovation and Strategy*, 1(1). Available at: https://doi.org/10.70818/pjbis.2024.v01i01.014.

Siami-Namini, S., Tavakoli, N. and Namin, A.S. (2019) “A Comparative Analysis of Forecasting Financial Time Series Using ARIMA, LSTM, and BiLSTM.” arXiv. Available at: https://doi.org/10.48550/arXiv.1911.09512.

Smith, K.A. and Gupta, J.N.D. (2000) “Neural networks in business: techniques and applications for the operations researcher,” *Computers & Operations Research*, 27(11–12), pp. 1023–1044. Available at: https://doi.org/10.1016/S0305-0548(99)00141-0.

Yamak, P.T., Yujian, L. and Gadosey, P.K. (2019) “A Comparison between ARIMA, LSTM, and GRU for Time Series Forecasting,” in *Proceedings of the 2019 2nd International Conference on Algorithms, Computing and Artificial Intelligence*. *ACAI 2019: 2019 2nd International Conference on Algorithms, Computing and Artificial Intelligence*, Sanya China: ACM, pp. 49–55. Available at: https://doi.org/10.1145/3377713.3377722.