Comparing Statistical and Machine Learning Models for Forecasting Milk Consumption in Ireland

MSc Research Project
Data Analytics

Utkarsh Arun Satpute
Student ID: x23135760

School of Computing
National College of Ireland

Supervisor: Teerath Kumar Menghwar

National College of Ireland

MSc Project Submission Sheet

School of Computing

Student Name:	Utkarsh Arun Satpute		
Student ID:	x23135760		
Programme:	Data Analytics	**Year:**	2024-25
Module:	Msc Research Project		
Supervisor:	Teerath Kumar Menghwar		
Submission Due Date:	11/08/2025		
Project Title:	Comparing Statistical and Machine Learning Models for Forecasting Milk Consumption in Ireland		

WordCount:7542 **Page Count** 21

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: Utkarsh Arun Satpute

Date: 11/08/2025

PLEASE READ THE FOLLOWING INSTRUCTIONS AND CHECKLIST

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

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

Office Use Only	
Signature:	
Date:	
Penalty Applied (if applicable):	

"Comparing Statistical and Machine Learning Models for Forecasting Milk Consumption in Ireland"

Utkarsh Arun Satpute
x23135760

Abstract

Accurate forecasting of dairy consumption is essential for effective supply chain management and policy planning for Ireland. From this research we will aims to explore and evaluate the performance of various time series forecasting models in predicting milk consumption trends across multiple categories, which include whole milk, skimmed milk, and semi-skimmed milk Irish dairy data. The goal is to identify the most accurate forecasting model for each category by comparing traditional statistical approaches with modern machine learning and deep learning techniques.

To acquire the result, we have implemented a data preprocessing technique in which it includes handling of missing values, stationarity, seasonal decomposition and testing of normalisation, and to evaluate the models we have used three type of evaluation metrics which are the (RMSE, MSE, R^2). We also did an comparative analysis of raw versus normalized data revealed notable differences in model behaviour. The results highlight varying strengths of each model depending on the milk category and data transformation was applied.

Among the tested models, SARIMA performed best for the whole milk and for the import and export dataset. While XGBoost outperformed others on skimmed milk. And one deep learning model that is LSTM achieved the strongest performance for the Semi – Skimmed milk. But other deep learning model Prophet underperformed on the irregular series. We also applied normalization which improved ML/DL models, but it reduced statistical accuracy of the models.

This study will contribute to a benchmark for forecasting techniques and offers practical insights for dairy producers and policymakers. The research concludes with a discussion on model selection, limitations and recommendations for future work, including deeper feature engineering and real-time predictive systems

1) Introduction

Irish diet is based mainly on Milk and dairy products, and it also contribute to agricultural economy, so it plays an important role in people's diet. Ireland is known for its grass-based dairy production system and is also one of the largest exporters of dairy products in Europe. Dairy accounts for is known portion of both household nutritional intake and the agri-food export sector, with products such as whole milk, cheese, and butter contributing to the country's GDP. As consumer preferences changes by time to time due to health trends, dietary shifts, and economic pressures there has been a noticeable variation in demand across different milk categories which include whole milk, skimmed milk, and semi-skimmed milk. These changing consumption patterns, when combined with seasonal, political, and environmental factors, introduce considerable uncertainty in supply chain operations. To make it more accurate the need for forecasting of dairy consumption become crucial for effective production planning, inventory management, and national food security strategies.

In the context of supply chain management, demand forecasting serves as a key decision-making tool. For the dairy sector, it informs procurement schedules, transportation logistics, processing capacities, and export commitments. Forecasting errors can lead to overproduction, spoilage, supply shortages, or unmet export contracts—all of which are detrimental to a perishable commodity like milk. The challenge, however, lies in the complex seasonality and non-linear patterns present in real-world milk consumption data. Traditional forecasting models such as the Autoregressive Integrated Moving Average (ARIMA), Seasonal ARIMA (SARIMA), and Holt-Winters exponential smoothing have been widely used for agricultural and time series forecasting due to their simplicity and interpretability. However, these models often struggle with non-linearities and multivariate dependencies in modern consumption data (Cesarini et al., 2024).

In recent years, deep learning (DL) techniques have gained traction in forecasting tasks due to their ability to learn complex patterns from large datasets. Long Short-Term Memory (LSTM) networks. And Prophet, developed by Facebook, offers a flexible additive model that handles seasonality and holidays well, making it a good choice for time series forecasting where interpretability is still desired (Iqbal et al., 2025; Mozher et al., 2024).

Despite these modern models, there is limited comparative research that benchmarks traditional statistical methods against modern Machine learning and Deep learning techniques on dairy consumption datasets, especially at the milk-type level. While Cesarini et al. (2024) conducted a comparative study of deep learning models for national milk forecasting but did not distinguish between types of milk. On other hand Mozher et al. (2024) focused on multi-retailer demand in the dairy sector and found that machine learning models like XGBoost outperformed traditional approaches but again did not explore disaggregated milk categories. Furthermore, little attention has been paid to the forecasting of import versus domestic milk intake which is an increasingly important distinction given Ireland's changing trade policies and global market dependencies.

This study seeks to address these gaps by developing a comparative forecasting framework that applies six distinct models like the ARIMA, SARIMA, Holt-Winters, Prophet, LSTM, and XGBoost on two datasets: (1) monthly milk consumption by category (Whole, Skimmed, Semi-Skimmed) from 1980 to 2024, and (2) yearly import vs. domestic milk intake volumes. In addition to comparing model performance across raw and normalized data, in which the study also analyses which models are more sensitive to data preprocessing, seasonality, and data sparsity.

Research Question

The research question guiding this study is:

"**Which forecasting models provide the highest accuracy for predicting Irish milk consumption trends across multiple categories, and how does normalization affect model performance, particularly in deep learning frameworks?**"

1.1 Research Objectives

To address this question, the following research objectives were defined:

1. To evaluate and compare the forecasting accuracy of ARIMA, SARIMA, Holt-Winters, Prophet, LSTM, and XGBoost on three milk categories: Whole, Skimmed, and Semi-Skimmed.
2. To investigate the impact of normalization on model performance, especially for LSTM and XGBoost.
3. To analyze import vs. domestic milk intake trends using the same modelling pipeline.
4. To visualize key findings and provide benchmarking insights for supply chain planners and policymakers.

1.2 Assumptions and Limitations

This study assumes that historical milk consumption patterns are reflective of future trends and that external shocks (e.g., pandemics, trade embargoes) are either embedded in the training data or negligible over the long term. A major limitation is the lack of demographic-level data (e.g., region, household size, age groups), which could offer finer granularity for forecasting. Also, the second dataset (Import vs. Domestic milk intake) contains fewer records (yearly data) compared to the monthly granularity of the first, affecting temporal resolution and model tuning. Furthermore, the study is limited by univariate time series which excludes the potential influentials like the external drivers' factors as the retail prices , consumer preferences and climate conditions due to this absence of factors restrict our study to the real work applicability

1.3 Contribution

The major contribution of this research lies in its comprehensive benchmarking of forecasting techniques on real, category-level dairy consumption data in Ireland. It is one of the few studies to include both classical and modern machine learning models, by applying them across multiple milk types, and evaluating performance under both raw and normalized conditions. The insights generated can help optimize dairy production and distribution pipelines, reduce forecasting error, and guide digital transformation strategies within the agri-food sector.

2) Related Work

2.1 Traditional Forecasting Models: ARIMA, SARIMA, and Holt-Winters

Traditional time series forecasting methods like the ARIMA (Auto-Regressive Integrated Moving Average), SARIMA (Seasonal ARIMA), and Holt-Winters Exponential Smoothing have been widely used for modeling agricultural and dairy-related data due to their mathematical transparency and ability to model seasonality trends. However, these models are much relied on the linearity, stationarity and consistent seasonality making them suitable for stable datasets and less effective for the irregular, volatile. Their univariate also limits real world applicability.
Musora et.al (2023) applied an ARIMA model specially SARIMA with (0,1,1) (0,1,1) for forecasting the demand of a perishable dairy drink. For this they used a real- world data in which they achieved a strong result of 90% in R^2 and, they achieved a low MAPE, showing that their model's where suitability for short-term seasonal forecasts. But they also highlight there where limitations in which model did not respond to sudden demand shifts which were caused by market volatility.

To broaden the perspective, a study by Panda et.al (2024) evaluated hybrid models for forecasting agricultural commodity prices, with combining CEEMDAN decomposition with time delay neural networks (TDNNs). The primary focus of the study was on hybrid deep learning, their results providing comparative insights into the limitations of traditional models like ARIMA and Holt-Winters when applied directly to volatile and nonlinear datasets. Their analysis showed that hybrid techniques significantly outperformed traditional models by reducing error metrics and improving adaptability to sudden demand fluctuations which is an issue often encountered in milk and dairy patterns. This supports the broader consensus that combining classical statistical methods with modern decomposition or machine learning frameworks can enhance robustness and predictive power. Further supporting this, Phumchusri and Sirimak (2024) conducted a comparative study involving SARIMA, Holt-Winters, and hybrid models with machine learning (ANN) for food product forecasting in Thailand. They concluded that Holt-Winters and SARIMA provided strong individual baselines, with Holt-Winters excelling in cases of consistent and demand patterns. However, the best performance was achieved when these traditional models were integrated into a hybrid architecture, significantly lowering MAPE across both validation and test datasets. This study demonstrated that evolving role of traditional models are not just the standalone predictors, but also an interpretable and effective components in modern forecasting ensembles.

Collectively, these studies validate the continued relevance of traditional statistical models for milk and dairy forecasting while also highlighting their limitations in capturing nonlinear, multivariate influences. This underscores the transition toward machine learning and hybrid approaches, which are explored in the following section.

2.2 Machine Learning Forecasting Models: LSTM, Prophet, and XGBoost

With the increasing demand in the dairy industry and the complexity for the production pattern, machine learning models have gained advance tools for forecasting. Unlike traditional statistical methods that assume linearity and stationarity, ML approaches offer flexibility in modeling irregular fluctuations, multivariate influences, and latent seasonal behaviors. Their ability to learn from data-driven patterns without rigid parametric structures makes them well-suited for agricultural scenarios, where external factors like feed quality, climate, and market dynamics often play a critical role in shaping production trends.

Alusyanti et al. (2023) explored the performance of LSTM and Prophet models in predicting daily goat milk production under limited-data scenarios in East Java. Their experiments showed that LSTM consistently outperformed Prophet, particularly when lag features were used in multivariate settings. The study emphasized the importance of LSTM's memory architecture in learning temporal dependencies even from small, irregular datasets, achieving an R^2 of up to 0.78 and MAPE below 4%.

In another relevant study, Kwarteng and Andreevich (2024) conducted a direct comparison of ARIMA, SARIMA, and Prophet models. Their findings revealed that while Prophet performed adequately on datasets with clear trends and daily frequency, it struggled with irregular series compared to SARIMA. This aligns with our own study, where Prophet underperformed on datasets like Skimmed and Semi-Skimmed Milk that lacked consistent seasonality. Their research emphasized the importance of choosing models based on data granularity and trend type rather than perceived model sophistication.

Adding to the growing body of deep learning applications in the dairy sector, a study from researchers at Fırat University by in Gür, Y.E. (2024) explored the performance of several machine learning and deep learning models—including LSTM, GRU, MLP, SVR, and KNN—for forecasting cow cheese production in Turkey. Although cheese is a derived dairy product, the authors argue that its production dynamics are closely linked to milk supply patterns and seasonal variations. Among the tested models, LSTM achieved the lowest RMSE and MAPE, demonstrating superior performance in capturing nonlinear trends and cyclical behaviors inherent in dairy time series. The study reinforces the adaptability and predictive power of LSTM in real-world agricultural forecasting contexts, especially when contrasted with traditional or shallow learning techniques.

2.3 Research Gaps and Comparison with This Work

While numerous studies have explored time series forecasting in the dairy sector, several limitations persist across the literature. Traditional models like ARIMA and SARIMA have been widely applied, but their inability to capture nonlinear dependencies, multivariate interactions, or abrupt trend shifts has been consistently highlighted. At the same time, many machines learning approaches, including LSTM and XGBoost, have demonstrated superior predictive accuracy, yet are often applied in isolation or without systematic comparison to statistical baselines. Moreover, few studies have evaluated model performance across multiple milk categories such as whole, skimmed, and semi-skimmed milk or they have explored the effects of normalization on forecasting accuracy.

This research will address these gaps by conducting a comparative analysis of six forecasting models: ARIMA, SARIMA, Holt-Winters, Prophet, LSTM, and XGBoost. Unlike prior works that only focus on a single milk type or ignoring data scaling, this study applies each model to multiple categories of milk consumption in Ireland and examines the influence of normalization techniques. By evaluating each method using consistent error metrics (RMSE, MSE, and R^2) and comparing results before and after normalization, the study will provide a more holistic understanding of forecasting performance in the context of evolving dairy consumption trends.

3) Research Methodology

This research aims to forecast milk consumption trends in Ireland by leveraging a combination of traditional statistical models and modern machine learning algorithms. The study is structured to support a comparative evaluation of model performance across multiple milk categories and data granularities. By applying multiple models—including ARIMA, SARIMA, Holt-Winters, Prophet, LSTM, and XGBoost—on two distinct datasets, this research ensures a balanced assessment of temporal forecasting methods across both raw and normalized data. The methodology is underpinned by a clearly defined workflow consisting of data acquisition, preprocessing, exploratory analysis, feature engineering, model training, and evaluation.

3.1 Research Design

In this study we have employed a research design that not only compare forecasting framework but also structure robust evaluation for predicting performance for different time series models. The core objective of the research is to forecast milk consumption trends in Ireland by applying and comparing both traditional statistical forecasting methods and modern machine learning techniques. This hybrid approach will ensure comprehensive model benchmarking under varied data assumptions for capturing both linear and nonlinear dynamics in the consumption patterns.

For this study we have used historical datasets which be used to train forecasting models and compare their generalization capability on unseen future data. This design supports replication, interpretability, and statistical evaluation, fulfilling academic requirements for transparency and scientific rigor. Following recommendations from the Research Methodology Design and Implementation guide, all steps in the research process are from data acquisition to model evaluation which are documented and reproducible.

Using both traditional and modern models stems from the distinct strengths they offer. Traditional models such as ARIMA, SARIMA, and Holt-Winters are interpretable and effective under stationary and seasonally consistent data. In contrast, models like LSTM and XGBoost offer the capacity to model complex, nonlinear dependencies and long-term memory in time series. The Prophet model bridges these paradigms with a modular, hybrid structure that can automatically adjust to changepoints and incorporate domain-specific holidays.

To validate these models, two publicly available datasets were selected from the Central Statistics Office (CSO) of Ireland:

- Dataset 1 includes monthly consumption data for Whole, Skimmed, and Semi-Skimmed Milk from 1980 to 2024.
- Dataset 2 captures annual milk intake categorized into Import and Domestic sources from 2000 to 2023

This diversity of models enables a rich comparative analysis across multiple axes such as interpretability, sensitivity to scale, ability to model seasonality, and predictive accuracy, thus offering a well-rounded assessment of forecasting effectiveness for policymaking, demand planning, and supply chain decisions in the Irish dairy sector.

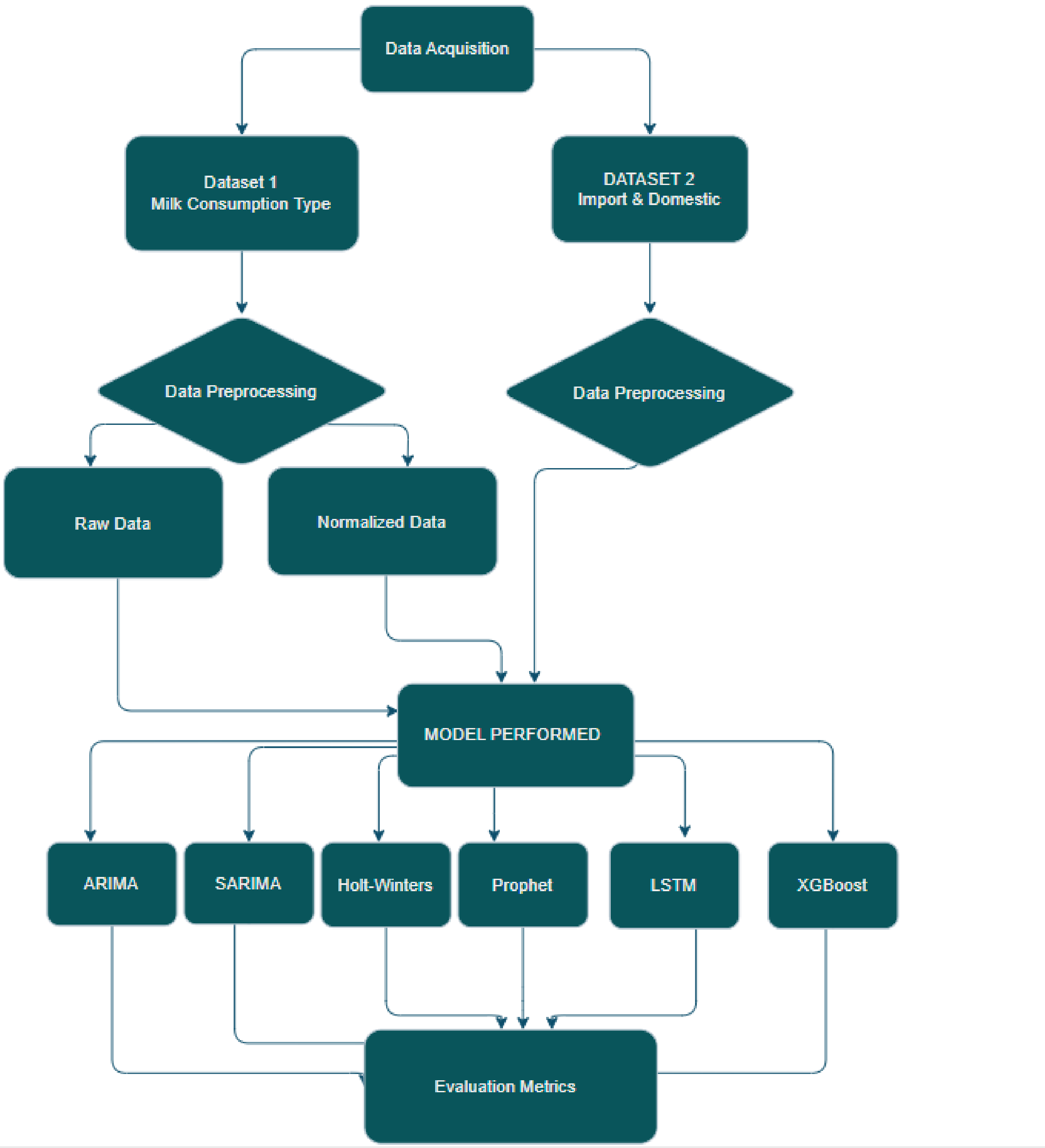

Figure 3.1: Research Methodology Workflow

3.2 Data Collection

The data for this research was collected from the Ireland's official site that is https://data.gov.ie/dataset/akm02-milk-sales-dairy-for-human-consumption, which is a reliable and publicly accessible source of national economic and agricultural statistics. Two distinct datasets were obtained to enable a comprehensive and comparative analysis of forecasting techniques across different levels of temporal resolution. These datasets represent both high-frequency monthly consumption data and low-frequency annual supply data, allowing for a robust evaluation of traditional statistical models and modern machine learning algorithms under diverse conditions.

The first dataset, referred to as Dataset 1, contains monthly milk consumption figures in litres from January 1980 to December 2024. It is segmented into four primary categories of milk: Whole Milk, Skimmed Milk, Semi-Skimmed Milk and All Milk. This dataset includes over 540 monthly records for each category, providing a rich time series ideal for detecting seasonal cycles, long-term consumption shifts, and public preference trends. The dataset was available in CSV format and consisted of time-based and categorical fields, namely the month of observation, the milk category, and the volume consumed. Owing to its high temporal granularity, Dataset 1 was used to train and evaluate all six forecasting models applied in this study: ARIMA, SARIMA, Holt-Winters, Prophet, LSTM, and XGBoost. Furthermore, both raw and normalized versions of this dataset were prepared to assess the impact of feature scaling on model performance, particularly for deep learning and ensemble methods.
The second dataset, Dataset 2, captures the total annual milk intake in Ireland between 2000 and 2023 it was downloaded from https://data.cso.ie/ , and is was categorized by source: Import or Domestic production. Although it is lower in frequency compared to Dataset 1, but Dataset 2 offers critical insights into the supply-side dynamics of the dairy sector over the last two decades. It includes 24 annual observations per category and was also obtained in structured CSV format. The dataset features three columns: year, milk source (import or domestic), and quantity in litres. Like Dataset 1, this dataset was used to train all six forecasting models, ensuring consistent methodological application across both high-frequency and low-frequency time series. The inclusion of Dataset 2 allowed the study to explore how forecasting models behave when applied to aggregated, low-volume datasets, and to evaluate their capacity to uncover meaningful trends over longer time horizons.
Both datasets were ethically sourced and are available for public use. The data is non-sensitive, anonymized, and requires no special permissions for academic use. The quality and completeness of the data made it suitable for direct integration into the forecasting pipeline without extensive transformation. Appropriate citations and acknowledgements to Ireland dataset site have been included throughout the research to ensure transparency and academic integrity.

3.3 Pre Processing and Cleaning

Effective data preprocessing is a foundational step in any time series forecasting pipeline. For this research, both Dataset 1 (monthly milk consumption) and Dataset 2 (annual import vs. domestic intake) were pre-processed using a systematic approach to ensure temporal consistency, structural completeness, and compatibility with various forecasting models. The goal was to produce clean, analysis-ready datasets suitable for both statistical and machine learning-based modelling techniques.
The initial phase of preprocessing involved parsing the time-based fields (Month or Year) into appropriate datetime formats and setting them as the index for each dataset. This step ensured proper time series ordering, which is essential for temporal analysis. Dataset 1 was verified to have a consistent monthly frequency across the full-time span from 1980 to 2024. Dataset 2, although lower in frequency, was similarly validated for yearly continuity between 2000 and 2023.
Missing values were minimal and occurred intermittently in Dataset 1. These gaps were handled using forward-fill and linear interpolation techniques, which are widely accepted for filling missing points in time series without disrupting the underlying trend or seasonality. This approach is supported by Zhou, Aryal, and Bouadjenek (2023), who emphasize that such imputation methods are effective when working with continuous data patterns that evolve gradually over time.

Once the missing values were resolved, then we perform a stationarity check using the Augmented Dickey-Fuller (ADF) test on both Dataset 1 and Dataset 2. This statistical test is essential for identifying the presence of unit roots in a time series, which can indicate non-stationarity. In cases where the ADF test confirmed non-stationary behaviour, first-order differencing was applied to remove trends and stabilize the mean. This step was particularly important before applying ARIMA and SARIMA models, which assume input data to be stationary for effective parameter estimation and accurate forecasting.

Further preprocessing steps included seasonal decomposition of Dataset 1 using additive models to extract components such as trend, seasonality, and residual noise. These decompositions helped guide model selection and informed the tuning of hyperparameters such as seasonal periods and smoothing factors. Dataset 2, being annual and shorter in length, was decomposed visually and statistically to detect any structural changes or trend shifts over time.

All preprocessing was conducted using Python's pandas, statsmodels, and scikit-learn libraries. The entire pipeline was version-controlled in Jupyter Notebooks to ensure reproducibility and traceability. These preprocessing steps laid the foundation for building robust forecasting models and performing fair comparative evaluations in subsequent phases of the study.

3.4 Normalization

Normalization is a crucial preprocessing technique that transforms numerical values into a common scale, typically between 0 and 1. This transformation is especially important when dealing with models that are sensitive to the magnitude of features or rely on gradient-based optimization. In time series forecasting, normalization ensures that variables with larger scales do not dominate model learning and also improves training stability, convergence speed, and generalization for certain types of models.

In this study, normalization was applied exclusively to Dataset 1, which contains monthly milk consumption data across Whole Milk, Skimmed Milk, and Semi-Skimmed Milk. The decision to normalize Dataset 1 was driven by its wide range of values across different milk categories and the longer time span, which introduced scale differences that could impact model performance. Dataset 2, being annual and lower in variation, was retained in its raw form to preserve its scale and macro-trend characteristics.

To comprehensively assess the impact of normalization, all six forecasting models—ARIMA, SARIMA, Holt-Winters, Prophet, LSTM, and XGBoost—were trained and evaluated on both the raw and normalized versions of Dataset 1. The MinMaxScaler from the scikit-learn library was used to transform the data into a [0, 1] range for each milk category. After forecasting, predictions were inverse transformed back to their original scale for evaluation using RMSE, MSE, and R^2. This choice was supported by a recent large-scale study by Lima and Souza (2023), which found that Min-Max normalization consistently enhanced predictive performance for deep learning models especially for LSTM by improving gradient stability and reducing convergence time in time series settings.

The results revealed that machine learning models, particularly LSTM and XGBoost, benefited significantly from normalization. These models rely on numerical stability during training and are affected by the relative scale of inputs. After normalization, they showed improved forecasting accuracy, with lower error metrics and better fit on unseen data. In contrast, traditional statistical models such as ARIMA, SARIMA, and Holt-Winters performed better on the raw data, where scale-dependent seasonal and trend components were preserved. Prophet, being partially scale-invariant, delivered stable performance across both normalized and raw inputs, although results were slightly better with raw data.

The normalization experiment offered valuable insight into how preprocessing choices influence model outcomes. It confirmed that not all models benefit equally from normalization, and that aligning preprocessing strategies with model architecture is essential for achieving optimal forecasting accuracy. By evaluating all models on both data versions, this study also reinforced methodological rigor, ensuring that the results were driven by empirical testing rather than assumptions. All normalization procedures and inverse transformations were implemented using Python's scikit-learn library, ensuring consistency, reproducibility, and proper scaling integrity.

3.5 Model Selection and Configuration

To ensure a well-rounded and domain-relevant forecasting framework, this study adopted six diverse models spanning both traditional statistical and modern machine learning paradigms: The selection was guided by both theoretical relevance and empirical evidence from recent literature, which ensure the models could collectively capture a variety of temporal behaviours like linear trends, seasonal components, changepoints, and nonlinear dependencies (Yaseen et al., 2024; Albeladi et al., 2023).

The statistical models like ARIMA, SARIMA, and Holt-Winters were chosen for their proven effectiveness in univariate time series analysis in which assumptions about stationarity and additive seasonality often hold. These models provide not only accurate forecasts in well-structured datasets but also transparency and interpretability, which are essential for policymaking in the dairy sector. For ARIMA and SARIMA, the optimal model parameters (p, d, q) and (P, D, Q, s) were selected based on minimizing the Akaike Information Criterion (AIC), but this was applied only for Dataset 1 under normalization. The Holt-Winters model was applied using additive seasonality without additional tuning, making it computationally efficient and easy to interpret.

Prophet was selected due to its modular framework that automatically handles seasonality's, holidays, and changepoints in time series data, making it especially useful in cases where underlying temporal patterns change over time. The model is also robust to missing data and outliers and was implemented using its default hyperparameters. Prophet served as a bridge between traditional and machine learning models due to its hybrid, rule-based structure (Yaseen et al., 2024).

For deep learning, the LSTM model was employed for its ability to learn long-range dependencies in sequential data, overcoming the limitations of traditional models in capturing nonlinear patterns and temporal shifts. A straightforward architecture was used comprising one or two LSTM layers followed by a dense output layer which is suitable for sequence-to-one forecasting. While no hyperparameter tuning was performed, the model leveraged normalized data (via MinMaxScaler) to stabilize gradient descent and improve convergence. This decision is aligned with findings from Albeladi et al. (2023), who noted that LSTM performance improved significantly with scale-adjusted inputs in time series tasks.

XGBoost, a powerful tree-based gradient boosting model, was included for its ability to model complex interactions without assuming linearity or seasonality. Although XGBoost is not natively designed for sequential data, lagged features were engineered to provide temporal context. It was implemented using default parameters to serve as a strong baseline. The inclusion of XGBoost was supported by prior studies such as Yaseen et al. (2024), which demonstrated its competitive performance alongside deep learning methods in time-sensitive domains like healthcare demand forecasting.

Overall, these six models were selected not only for their individual strengths but also for their collective ability to cover the forecasting spectrum—from high interpretability and simplicity to high flexibility and predictive power. This design enabled robust comparative analysis and allowed insights into how different modelling paradigms perform under varying preprocessing conditions, data frequencies, and milk categories.

4) Results and Analysis

4.1 Forecasting Results on Whole Milk (Raw Data)

For the forecasting of whole milk, we have applied all the six models on this data. For this we have split the data into train and test. For training we have used from the January 1980 to December 2020, and we will test that from the January 2021 to July 2024. To get output we have used traditional and modern machine learning models. And to evaluate them perform we will be using three evaluation metrics and that RMSE, MSE and R^2 score

After applying these six models in that SARIMA performed the best scoring the RMSE of 1.42, MSE as 0.99 and R^2 0.640 and the second model which performed good was the Holt-Winter with the RMSE, MSE and R^2 of 1.88, 1.45 and 0.3839. and the third model was the XGBOOST with 1.99 ,1.58 and 0.3095 as the performance.

While on other hand, Models like the LSTM and ARIMA showed weak predictive ability and the Prophet performed worst with the result getting to negative of -1.1942 R^2 score.

MODEL	RMSE	MSE	R^2 SCORE
SARIMA	1.42	0.99	0.6460
Holt-Winters	1.88	1.45	0.3839
XGBOOST	1.99	1.58	0.3095
LSTM	1.94	1.54	0.1741
ARIMA	2.39	1.96	0.0068
PROPHET	3.55	3.12	-1.1942

Table 4.1: Model Evaluation Metrics for Whole Milk Forecasting

To understand clearly the performance of model for the whole milk, I have illustrated table 4.1. In which we can see what the ranking of the models are.

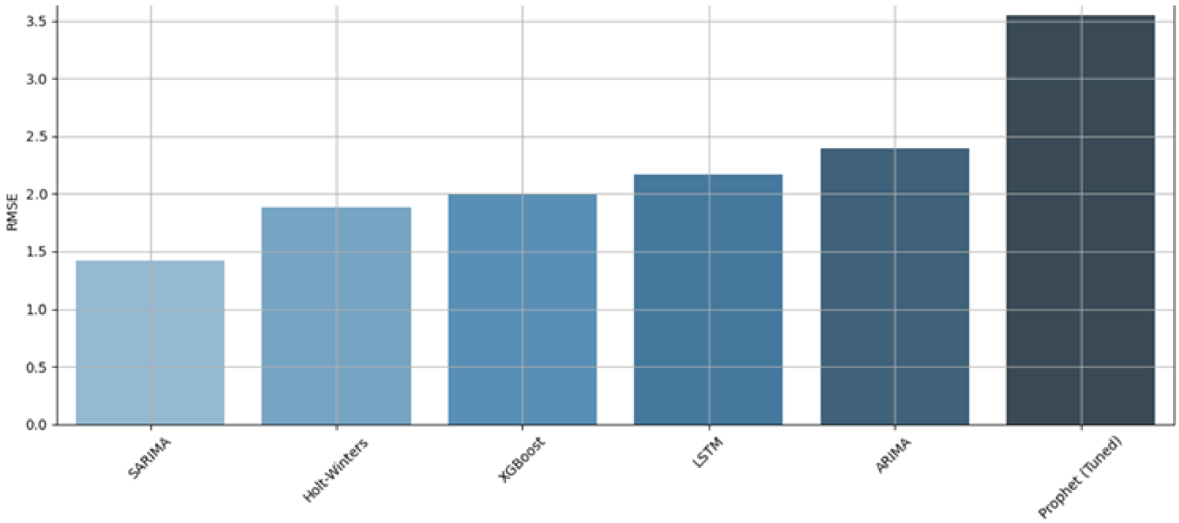

Figure 4.1: Model Performance Comparison on Whole Milk (Raw Data)

Overall, the SARIMA model emerged as the most effective approach for forecasting Whole Milk consumption in this context. Its strong performance demonstrates the importance of explicitly modeling seasonal differencing and autoregressive structures in time series with clear periodic patterns. While machine learning models like XGBoost showed potential in handling irregular patterns, traditional statistical methods remained more reliable for this specific dataset.

4.2 Forecasting Results on Whole Milk (Normalized Data)

In this also SARIMA model gave the best result with the 0.0586 RMSE, 0.0406 MSE and 0.6461 R^2. The holt-winter remained at the same position before applying and after applying normalization, but result changed little bit securing (0.0773 RMSE, 0.0595 MSE, 0.3843 R^2). But due to normalization the result for the LSTM improved and it became the third good performer with the RMSE, MSE and R^2 (0.0796, 0.0637, 0.1819). XGBoost delivered a moderate performance with RMSE = 0.0822, MSE = 0.0655, and R^2 = 0.3040. In contrast, ARIMA performed poorly (RMSE = 0.1044, MSE = 0.0849, R^2 = –0.1235), and Prophet was the weakest model, producing RMSE = 0.1377, MSE = 0.1208, and a negative R^2 = –0.9525.

MODEL	RMSE	MSE	R^2 SCORE
SARIMA	0.0586	0.0406	0.6461
Holt-Winters	0.0773	0.0595	0.3843
LSTM	0.0796	0.0637	0.1819
XGBOOST	0.0822	0.0655	0.3040
ARIMA	0.1044	0.0849	- 0.1235
PROPHET	0.1377	0.1208	-0.9525

Table 4.2: Ranked Model Evaluation Metrics for Whole Milk Forecasting (Normalized Data)

The ranked performance of all models is summarized in Table 4.2, based on RMSE, MSE, and R^2 score. A visual comparison of RMSE values between raw and normalized versions is presented in Figure 4.2

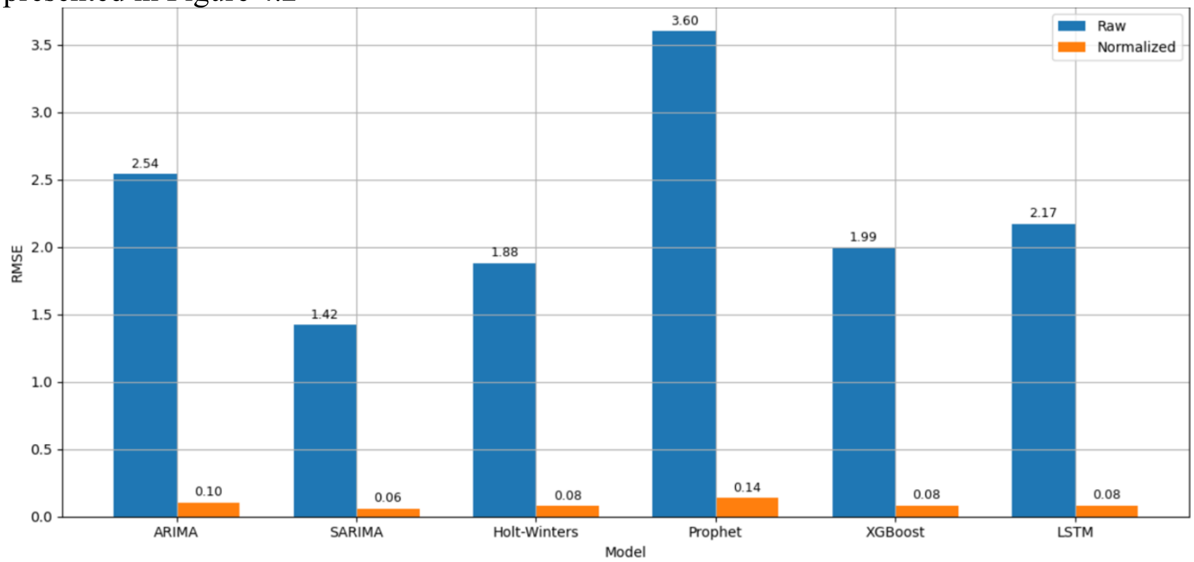

Figure 4.2: **RMSE Comparison Between Raw and Normalized Forecasts for Whole Milk Consumption (1980–2024).**

While normalization compresses the scale of the data and helps with model convergence especially for neural networks like LSTM, but it does not necessarily improve accuracy across all models. In this study, SARIMA continued to outperform others even without the need for scaling, reaffirming its effectiveness in modelling seasonal dairy consumption patterns.

4.3 Forecasting Results on Skimmed Milk

Forecasting for the Skimmed Milk consumption was a bit challenging because of its irregular seasonal fluctuations and lower overall volume data compared to Whole Milk. When applying six forecasting models like the ARIMA, SARIMA, Holt-Winters, Prophet, LSTM, and XGBoost we saw a clear disparity emerged between traditional and machine learning-based approaches.

Traditional statistical models struggled to generalized well on the test data. But ARIMA marginally better among the three, recorded an RMSE of 0.443 and MSE of 0.421, with an R^2 score of -1.388. While on other hand, SARIMA and Holt-Winters performed even worse, reflecting poor seasonal adaptation, particularly for low-volume series with weak trends. Both models yielded RMSE values above 0.49 and highly negative R^2 scores (–1.87 and –1.95 respectively), indicating severe model underfit and low predictive utility.

Moreover, modern approaches were better suited to the subtle patterns in Skimmed Milk sales. Prophet showed some seasonal alignment, reducing RMSE to 0.378 and MSE to 0.355, although it still returned a negative R^2 score (–0.745). LSTM captured temporal dependencies more effectively, with achieving a notably lower RMSE of 0.296 and MSE of 0.258, with a nearly neutral R^2 (–0.067), indicating partial improvement in variance explanation.

XGBoost emerged as the top performer for Skimmed Milk forecasting. With an RMSE of just 0.222 and MSE of 0.182, it not only achieved the lowest error rates but also the only positive R^2 value (0.401), reflecting meaningful predictive power. This suggests that tree-based models may offer better performance on datasets with limited seasonality but complex lag-based dependencies.

MODEL	RMSE	MSE	R² SCORE
XGBOOST	0.222	0.182	0.401
LSTM	0.296	0.258	-0.067
PROPHET	0.378	0.355	-0.745
ARIMA	0.443	0.421	-1.388
SARIMA	0.490	0.460	-1.870
HOLT-WINTERS	0.492	0.467	-1.947

Table 4.3 – Ranked Model Evaluation Metrics for Skimmed Milk Forecasting

Table 4.3 summarizes the ranked evaluation metrics for all six models. Figure 4.3 presents a comparative visualization based on RMSE, MSE, and R^2 scores, this clearly highlights the XGBoost and LSTM models as the top-performing models, while traditional statistical methods consistently underperformed.

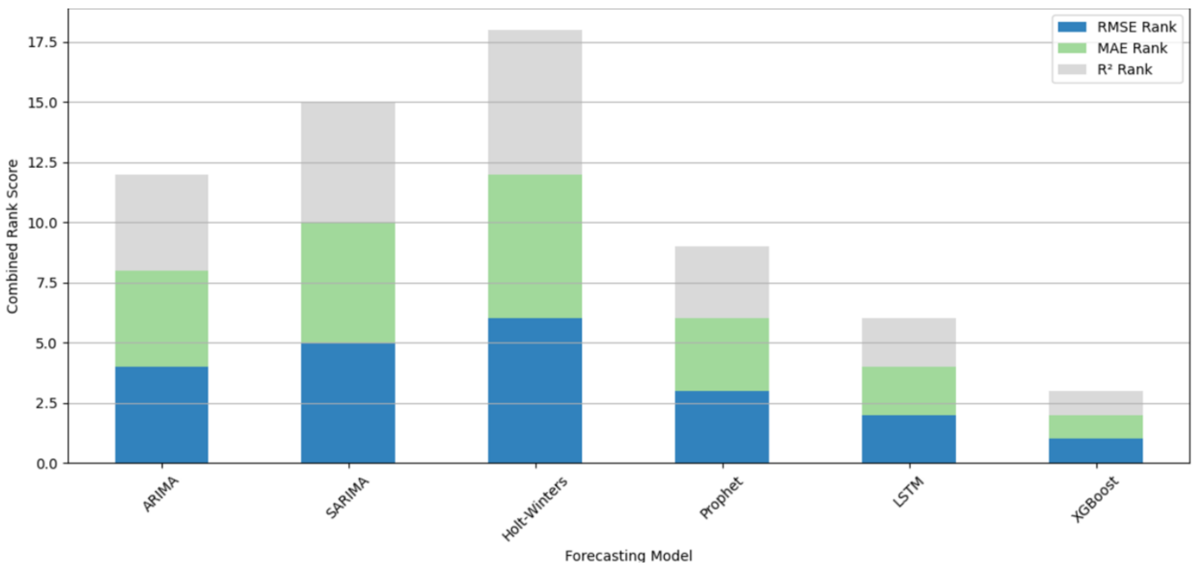

Figure 4.3: Comparison of Forecasting Model Performance on Skimmed Milk Consumption

4.4. Forecasting Result on Semi – Skimmed Milk Consumption

For semi-skimmed milk consumption, the traditional models like the ARIMA gave comparatively better results, recording an RMSE of 0.8756 and MSE of 0.6696, yet it still failed to explain much variance with an R^2 of -0.4467. While SARIMA and Holt-Winters models performed significantly worse, producing RMSE values of 1.9936 and 2.0425, respectively, and extremely negative R^2 scores (−6.5 and −6.87), indicating an inability to effectively model the seasonal and trend components in the data. The Prophet model showed the poorest alignment with actual values, with an RMSE of 2.53 and R^2 of −11.11, highlighting severe limitations when applied to this dataset.
On other hand, the Long Short-Term Memory (LSTM) model stood out, delivering the most accurate forecasts with an RMSE of 0.66 and MSE of 0.55, alongside a marginally positive R^2 score of 0.060 suggesting that it could partially capture the temporal dynamics and nonlinear patterns in the series. XGBoost also demonstrated competitive performance (RMSE: 0.77, MSE: 0.66), although its R^2 remained slightly negative at −0.280.

MODEL	RMSE	MSE	R² SCORE
LSTM	0.660	0.550	0.060
XGBoost	0.770	0.660	−0.280
ARIMA	0.8756	0.6696	−0.4467
SARIMA	1.9936	1.840	−6.500
HOLT-WINTERS	2.0425	1.7416	−6.870
PROPHET	2.530	2.230	−11.110

Table 4.4 – Ranked Model Evaluation Metrics for Semi-Skimmed Milk Forecasting

The comparative performance of all six forecasting models is summarized in Table 4.4, which ranks them based on RMSE, with MSE and R^2 scores provided for additional context. To further aid interpretation, a stacked bar chart is presented in Figure 4.4, which offer a visual comparison across all three-evaluation metrics. The figure clearly demonstrates that the LSTM model consistently outperformed the others, followed by XGBoost and ARIMA, while traditional statistical models (SARIMA, Holt-Winters, and Prophet) exhibited significantly lower predictive accuracy. These findings reinforce the effectiveness of deep learning techniques in capturing complex and non-linear consumption trends, which traditional methods often fail to model accurately.

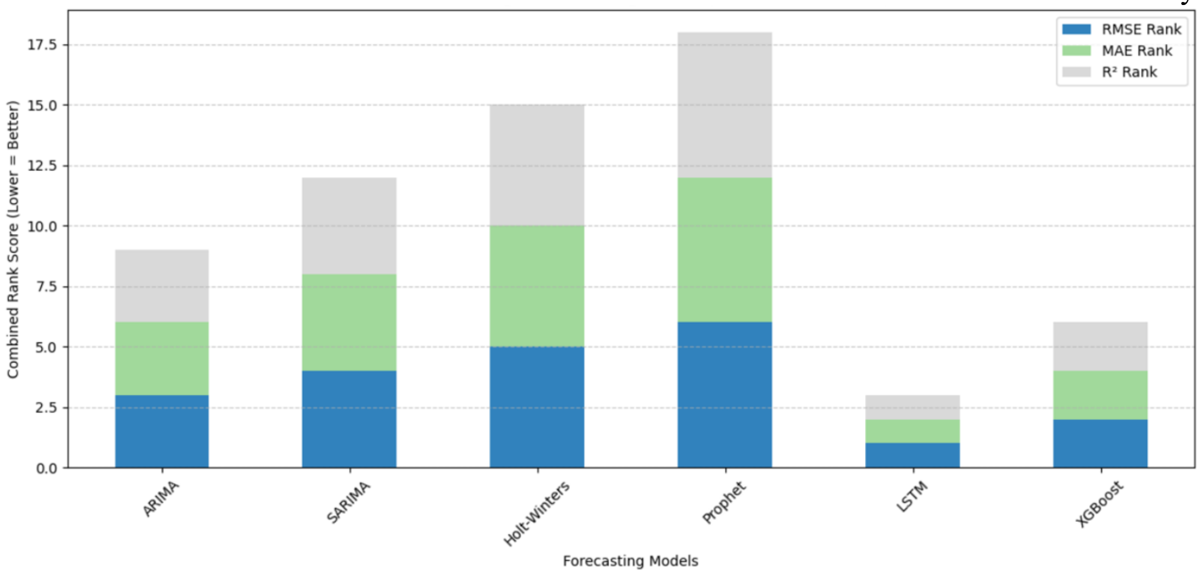

Figure 4.4: Stack Ranking of Forecasting Models for Semi – Skimmed Milk

4.5 Forecasting Result on the Import & Export

For the Import & Export, SARIMA and Holt-Winter performed the strongest and consistent result as compared to other models they acquired RMSE, MSE and R^2 Score same (38.45, 30.87 and 0.9841). The XGBOOST achieved the third position with an RMSE of 39.50, MSE of 31.04, and a strong R^2 Score of 0.9832 showing that the gradient boost can rival with traditional model if they are tunned properly. ARIMA and LSTM models performed the worst, while ARIMA failed to capture trend and seasonality and LSTM suggested overfitting due to its limited sequential patterns. Both these model's R^2 score was in the negative.
A detailed comparison of all six models is presented in Table 4.5, ranked by RMSE for interpretability across all evaluation metrics.

MODEL	RMSE	MSE	R^2 SCORE
SARIMA	38.45	30.87	0.9841
Holt-Winters	38.45	30.87	0.9841
XGBOOST	39.50	31.04	0.9832
PROPHET	127.58	84.99	0.8247
ARIMA	874.01	768.41	−7.2253
LSTM	804.34	745.65	−8,936,662.0

Table 4.5: Ranked Model Evaluation Metrics for Domestic Milk Intake Forecasting

To complement the tabular evaluation, a bar chart is presented in Figure 4.5, illustrating the relative performance of each model based on RMSE. The visual clearly reinforces the ranking shown in the table, confirming that SARIMA, Holt-Winters, and XGBoost are the top-performing models for forecasting Domestic milk intake, while ARIMA and LSTM trail significantly behind.

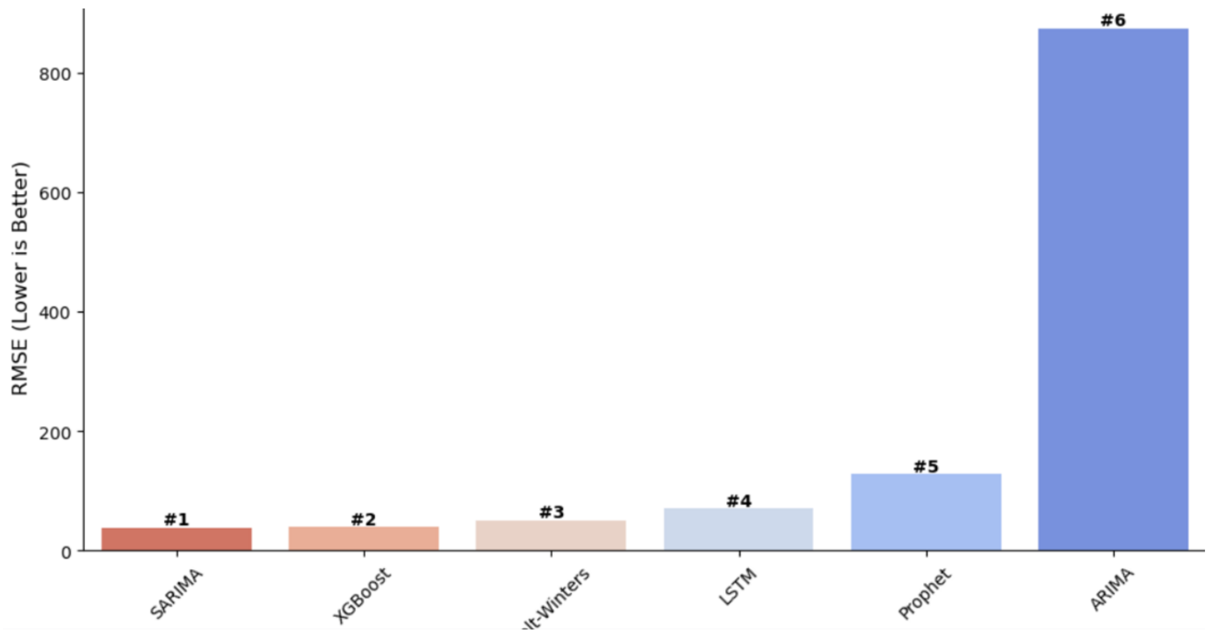

Figure 4.5: Model Performance Ranking by RMSE

5) Discussion

5.1 Evaluation of Forecasting Models

The results of traditional and modern models show consistent performance across all the datasets and has also mentioned from 4.1 – 4.5. For the whole milk with and without normalisation, SARIMA model delivers the lowest error score resulting the best performer and effective for modeling seasonal demand. While Holt winter and XGBoost ranked the second and third for the raw whole milk data and LSTM secure the third ranked. While seeing this result ranking it aligns with the Musora et al. (2023) , which report that SARIMA performance in short-term is stronger.

The forecasting for the skimmed milk was a bit struggling was difficult for traditional model due to the irregular and low volume pattern present in the data. ARIMA, SARIMA and Holt-winter all produced high RMSE values and strongly negative R^2 due to the weakness in handling volatile demand. Whereas XGBoost performed good result with RMSE = 0.222, MSE = 0.182, R^2 = 0.401, which captured nonlinear dependencies more effectively. While LSTM ranked second, which showed improvement temporal in learning compared to Prophet.

For the Semi-Skimmed Milk. LSTM model achieved the best result securing the first rank with outperforming all the models. The XGBoost and ARIMA was the next best models. While all the other models where underfitting with large error and high negative R^2 Score. This confirms Saha et al. (2024), who noted that LSTM memory architecture allows it to capture nonlinear temporal dynamics but only when it had sufficient sequence depth and data normalization is applied.

And lastly for the Import & Export, SARIMA and Holt Winter achieved the best performance with the XGBoost very close. While ARIMA and LSTM underperformed due to limited yearly data and Prophet again failed to capture trend irregularities.

Overall, the result confirms that choosing the model should be guided by dataset characteristics. In our case, Statistical model excels where seasonality dominant, for the machine learning models like XGBoost adapted well irregular pattern and LSTM showed the potential for nonlinear dynamics when is data is sufficient. Normalization improved neural network performance, but it reduced statical model accuracy

5.2 Model Limitations and Design Critique

While the comparative results across datasets demonstrated clear model rankings in terms of performance, a critical evaluation of the model designs and their limitations reveal several areas where improvements could be made.

LSTM model suffered a major challenge of poor generalisation ability on the low dataset. As we know deep learning models are known for their capturing long-term temporal dependencies, but their performance can decline with the sparse data and insufficient tuning. In our case, the absences of hyperparameter optimisation and data augmentation this two led to an unstable training behaviour and negative R^2 values. This shows that the findings from Saha et al. (2024), who showed from his paper that LSTM can collapse or overfit when the sequence length is inadequate or the augmentation is absent.

The second limitation was for the traditional models like SARIMA and Holt-Winters, which performed strongly for the whole milk and import and export milk, in which they are clearly seasonal and stable in long-term trends. Their seasonal differencing and smoothing mechanisms allowed them to capture recurring consumption cycles, resulting to get high R^2 values and Low error metrics. These features make them a strong and reliable forecasting for high volume categories with predictable patterns.

However, their assumption of linearity and stable seasonality proved wrong for the Skimmed and Semi-skimmed milk in which they had a volatility, irregular fluctuations and lower data volumes. In which they were unable to adapt to sudden shifts leading to large RMSE and strong negative R^2 score. This limitation reflects a broader weakness of traditional statistical approaches: while interpretable and computationally efficient, they lack the flexibility required to handle irregular, noisy, or low-signal time series.

The model that performed strongly was the XGboost which has few notable limitations. Its performance was dependent heavily on careful feature engineering and hyperparameter tuning, without that it risks overlooking subtle temporal dependencies. In the multi-step forecasts, boosting methods can also accumulate errors, reducing stability over longer horizons. As Casolaro et al. (2022) note that, even high performing ensemble approaches require the design choice to be ensured robustness in complex time series applications.

Another important limitation of this study was the reliance on univariate time series data. We exclude external drivers such as pricing, demographics or climate conditions, which the model failed to account for the real-world factors that strongly influences milk consumptions. This resulting that forecasts are statistically valid but their applicability to decision making in a dynamic market environment is reduced.

The study also highlights the mixed impact of normalization. In the performance of LSTM improved because of scaling and stabilizing training and accelerating convergences, yet it slightly reduced the accuracy of statistical models like the SARIMA and Holt winters. From this finding we can say that preprocessing strategies should be made according to the models.

Additionally, the limited forecasting horizon (restricted to 24 months) may have constrained our ability to assess the models' true long-term predictive power. Hybrid and ensemble models, such as those integrating ARIMA with XGBoost or LSTM with wavelet transforms, have been shown to significantly outperform individual models in complex forecasting scenarios. As highlighted by Udristioiu et al. (2024), hybrid frameworks combining statistical and machine learning techniques can enhance robustness, especially in volatile or noisy domains like air quality forecasting. Similar architectures could be explored in future dairy consumption research for improved accuracy and adaptability.

Another point of critique lies in the lack of external features. All models were trained purely on univariate time series data, excluding potentially impactful variables such as price changes, population demographics, or seasonal promotions. While this was intentional to

isolate temporal modelling performance, it inherently limited the explanatory power of all models.
In conclusion, while SARIMA, Holt-Winters, and XGBoost emerged as effective forecasting tools within this study, their limitations—along with those of Prophet and LSTM—must be acknowledged. Incorporating more advanced architectures, multivariate features, and augmentation techniques may bridge current performance gaps and offer more reliable insights for dairy consumption forecasting in the future.

5.3 Business Implications

The findings from this research have several important implications for both dairy industry stakeholders and policy makers involved in supply chain planning, demand forecasting, and inventory management. Accurate milk consumption forecasts can directly influence procurement decisions, storage optimization, and pricing strategies across various milk categories.
Firstly, from the study we can say the models like ARIMA, SARIMA and Holt Winters can be considered as the stable models for the high-volume data. So these models can be used for the dairy cooperatives and processing companies which could prevent overstocking and reduce spoilage. For instance, SARIMA's strong performance on Whole Milk data indicates its suitability for long-term forecasting in high-volume segments with consistent consumption trends.
Secondly, machine learning models like XGBoost showed superior performance in forecasting low-volume or irregular series, such as Skimmed Milk. These could be used for the smaller dairy producers or niche retailers, who often struggle with unpredictable demand for specialty milk types, they can benefit from adopting advanced, non-linear models to better anticipate sales and streamline logistics. The ability of XGBoost to capture non-obvious patterns is particularly valuable in environments with consumer behaviour shifts influenced by health trends or market disruptions.
At the same time, LSTM emerged as the best model for the Semi-skimmed milk showing the potential capturing complex temporal dynamics.
Moreover, the integration of such forecasting models into digital dashboards or ERP systems could support real-time decision-making. Dairy companies using tools like Power BI, Tableau, or Salesforce can embed these predictive outputs to monitor supply-demand alignment dynamically, allowing for quicker adjustments in distribution, promotion, and sourcing.
From a policy standpoint, the accuracy of milk intake predictions can aid national food agencies and cooperatives in ensuring food security and sustainable production planning. Forecasts of import and export milk intake, as modelled in Dataset 2, offer insights into import dependency and can help to optimize trade strategies or subsidy allocation in response to projected deficits or surpluses.
Finally, this research reinforces the need for data-driven decision-making in agriculture, especially in the face of global volatility in consumer demand and environmental constraints. Forecasting not only helps cut operational costs but also promotes sustainability by minimizing overproduction and reducing dairy waste across the value chain.

6) Conclusion and Future Work

In this study we have applied various time series models like the ARIMA, SARIMA, Holt Winter, Prophet, LSTM and XGboost. And these models were evaluated with the RMSE, MSE and R^2 to predict the milk consumption pattern in Ireland with the four datasets: Whole, Semi – skimmed milk, skimmed, and import and export dataset.
The results shows that SARIMA was the best performer for the whole milk before and after the normalisation, these result shows that its strength in modelling stable and seasonal demand. While for the skimmed milk XGBoost forecasting was best, in which it captured the irregular and low volume demand pattern more effectively than other models. And for the Semi-Skimmed milk LSTM achieved the best performance which shows the deep learning capturing capability for nonlinear and complex temporal, dynamics, with sensitivity data volume and tuning. Finally, for the import and export data the forecast by SARIMA and Holt-Winters was done accurately, while XGBoost result was very close. On other hand, Prophet consistently underperformed across all the datasets, producing high error values and strongly negative R^2.
The results demonstrate that traditional models like SARIMA and Holt-Winters performed best on datasets with strong seasonality, such as Whole Milk and Domestic Milk. XGBoost consistently emerged as a top performer, particularly in irregular or low-volume datasets like Skimmed Milk, due to its ability to model non-linear dependencies. Conversely, LSTM showed poor performance when data was sparse, especially for Domestic Milk, reinforcing the importance of data volume and stability in deep learning applications. Prophet, while interpretable, was found to underperform in most scenarios.
These findings validate the hypothesis that model selection must align with dataset characteristics—seasonality, trend stability, and data volume—rather than relying solely on model complexity. The study also underscores the practical value of combining traditional and machine learning approaches, depending on the forecasting context.
However, this research is not without limitations. LSTM underperformance may have been influenced by suboptimal hyperparameter tuning or inadequate sequence length, and XGBoost performance could be further optimized with advanced feature engineering. Additionally, external variables such as weather, economic trends, or promotional events were not included, which may affect real-world consumption.
From these findings we can say that the model selection should be according to the datasets. Traditional model has a good result in stable and seasonal series which machine learning models like the XGBoost are good adapter to irregular patterns and Deep learning models like LSTM provide value in mixed or semi-regular datasets when the sufficient data is available.

Future Work

For our study the future work research should be involvement of the hybrid forecasting models. In which they should combine both the traditional and modern machine learning models. Another addition was to include the data of retail prices, population demographics and environmental condition.
From a methodological perspective, Transfer Learning or Transformer-based architectures could be examined to address LSTM's data dependency issues, especially in low-volume series. Additionally, anomaly detection frameworks can be integrated to identify sudden shifts in consumption caused by policy changes or unexpected events like pandemics.
On a commercial level, this research can be extended into real-time forecasting systems integrated into supply chain dashboards for dairy producers and distributors. Automating inventory decisions, optimizing procurement cycles, and responding to demand volatility with predictive accuracy could unlock significant cost savings and reduce food waste.

References

Cesarini, L., Rui Gonçalves, Martina, M., Romão, X., Monteleone, B., Pereira, F.L. and Figueiredo, R. (2024). Comparison of deep learning models for milk production forecasting at national scale. *Computers and Electronics in Agriculture*, 221, pp.108933–108933. doi:https://doi.org/10.1016/j.compag.2024.108933.

Mozher, A., Fquiyeh, S., Lmariouh, J. and Soulhi, A. (2025). Demand Forecasting in Multi-Retailer Supply Chains Within the Dairy Sector. *The 1st International Conference on Smart Management in Industrial and Logistics Engineering (SMILE 2025)*, [online] p.48. doi:https://doi.org/10.3390/engproc2025097048.

Musora, T., Chazuka, Z., Jaison, A., Mapurisa, J. and Kamusha, J. (2023). Demand Forecasting of a Perishable Dairy Drink: An ARIMA Approach. *International Journal of Data Mining & Knowledge Management Process*, 13(1/2), pp.35–49. doi:https://doi.org/10.5121/ijdkp.2023.13203.

Musora, T., Chazuka, Z., Jaison, A., Mapurisa, J. and Kamusha, J. (2023). Demand Forecasting of a Perishable Dairy Drink: An ARIMA Approach. International Journal of Data Mining & Knowledge Management Process, [online] 13(1/2), pp.35–49. doi:https://doi.org/10.5121/ijdkp.2023.13203.

Pandit, P., Sagar, A., Ghose, B., Paul, M., Ozgur Kisi, Dinesh Kumar Vishwakarma, Mansour, L. and Yadav, K.K. (2024). Hybrid modeling approaches for agricultural commodity prices using CEEMDAN and time delay neural networks. Scientific Reports, [online] 14(1). doi:https://doi.org/10.1038/s41598-024-74503-4.

Phumchusri, N. and Sirimak, W. (2024). Time Series and Machine Learning Hybrid Models for Food Condiment Demand Forecasting: A Case Study in Thailand. International Journal of Machine Learning, 14(1), p.1. doi:https://doi.org/10.18178/ijml.2024.14.1.1149.

Gür, Y.E. (2024). Innovation in the dairy industry: forecasting cow cheese production with machine learning and deep learning models. International Journal of Agriculture Environment and Food Sciences, 8(2), pp.327–346. doi:https://doi.org/10.31015/jaefs.2024.2.9.

Kwarteng, S. and Poguda Andreevich (2024). Comparative Analysis of ARIMA, SARIMA and Prophet Model in Forecasting. Research & Development, 5(4), pp.110–120. doi:https://doi.org/10.11648/j.rd.20240504.13.

Primawati, A. and Astuti, D. (n.d.). Performance LSTM and Prophet for Prediction Time Series with Limited Data: Case Study of Daily Goat Milk Production 1. doi:https://doi.org/10.1109/ICON-SONICS59898.2023.10435067.

Zhou, Y., Aryal, S. and Bouadjenek, M. (n.d.). *A Comprehensive Review of Handling Missing Data: Exploring Special Missing Mechanisms*.

Lima, F. and Souza, V. (2023). A Large Comparison of Normalization Methods on Time

Series. *Big Data Research*, 34, p.100407. doi:https://doi.org/10.1016/j.bdr.2023.100407.

Albeladi, K., Zafar, B. and Mueen, A. (2023). Time Series Forecasting using LSTM and ARIMA. International Journal of Advanced Computer Science and Applications, 14(1). doi:https://doi.org/10.14569/ijacsa.2023.0140133.

Ripatti, V. (2025). Forecasting Prescription Medication Utilization: A Comparative Study of SARIMA, Prophet, XGBoost and LSTM Models. [online] Aalto.fi. Available at: https://aaltodoc.aalto.fi/items/56bf87d5-925d-4ee4-a437-239bb3c113cd [Accessed 30 Jul. 2025].

Hall, T. and Rasheed, K. (2025). A Survey of Machine Learning Methods for Time Series Prediction. *Applied Sciences*, [online] 15(11), p.5957. doi:https://doi.org/10.3390/app15115957.

Li, J., Lin, B., Wang, P., Chen, Y., Zeng, X., Liu, X. and Chen, R. (2024). A Hierarchical RF-XGBoost Model for Short-Cycle Agricultural Product Sales Forecasting. *Foods*, 13(18), pp.2936–2936. doi:https://doi.org/10.3390/foods13182936.

Rygh, T., Vaage, C., Westgaard, S. and Petter (2025). Inflation Forecasting: LSTM Networks vs. Traditional Models for Accurate Predictions. *Journal of risk and financial management*, [online] 18(7), pp.365–365. doi:https://doi.org/10.3390/jrfm18070365.

Zhao, T., Chen, G., Sujitta Suraphee, Tossapol Phoophiwfa and Piyapatr Busababodhin (2025). A hybrid TCN-XGBoost model for agricultural product market price forecasting. *PLoS ONE*, [online] 20(5), pp.e0322496–e0322496. doi:https://doi.org/10.1371/journal.pone.0322496.

Casolaro, A., Capone, V., Iannuzzo, G. and Camastra, F. (2023). information Deep Learning for Time Series Forecasting: Advances and Open Problems. doi:https://doi.org/10.3390/info14110598.

Mihaela Tinca Udriştioiu, Y. El Mghouchi and Hasan Yıldızhan (2023). Prediction, modelling, and forecasting of PM and AQI using hybrid machine learning. *Journal of Cleaner Production*, 421, pp.138496–138496. doi:https://doi.org/10.1016/j.jclepro.2023.138496.

Saha, S., Haque, A. and Sidebottom, G. (n.d.). Overcoming data limitations in internet traffic forecasting: LSTM models with transfer learning and wavelet augmentation. doi:https://doi.org/10.1016/j.comcom.2025.108280.