National College of Ireland

Group Project Title: Exploring Time Series Algorithms to Forecast Sales Targets in Retail Sector

Nachiket Anil Mehendale

V1.0

1. Introduction:

As a part of the Programming of AI – TABA (Sem1, MSCAI1, Batch Sep 2024), we as a part of the Group C, decided to do a research project in “Time-Series in the retail domain. The title of the group project is “Exploring Time Series Algorithms to Forecast Sales Targets in Retail Sector”. We found the Time-Series is an important topic in Machine Learning and it is used in almost every industry for the forecasting purpose.

2. Objectives:

To explore various techniques in Time-series was the primary objective and we chose a “Retail Sector” because plenty of open-source retail datasets are available online and so dataset availability was no longer an issue. Each member chose a different dataset and studied time-series analysis techniques, visualizations and algorithms. I chose a public dataset called “Monthly Milk Sales” published by Central Statistics Department of Ireland. The objective was to understand the time-series modelling techniques and using that forecast the milk sales. Forecasting is important to make informed decisions not just for the consumers but also for the dairy companies, farmers and policymakers. Forecasting reduces wastage, optimize production process and keep the supply stable in the local market.

3. Project Activities:

At a group level:

Research Topic Selection
Proposed the topic of Time-Series domain and coordinated with the rest of the team members.

Proactive coordination
Kicked off the discussion on the topic selection, and provided my inputs towards completing the project and at the same remained open to other viewpoints.

Report Creation
Initiated with the project report and distributed amongst the team with clear inputs on what and where to write their respective piece of code.

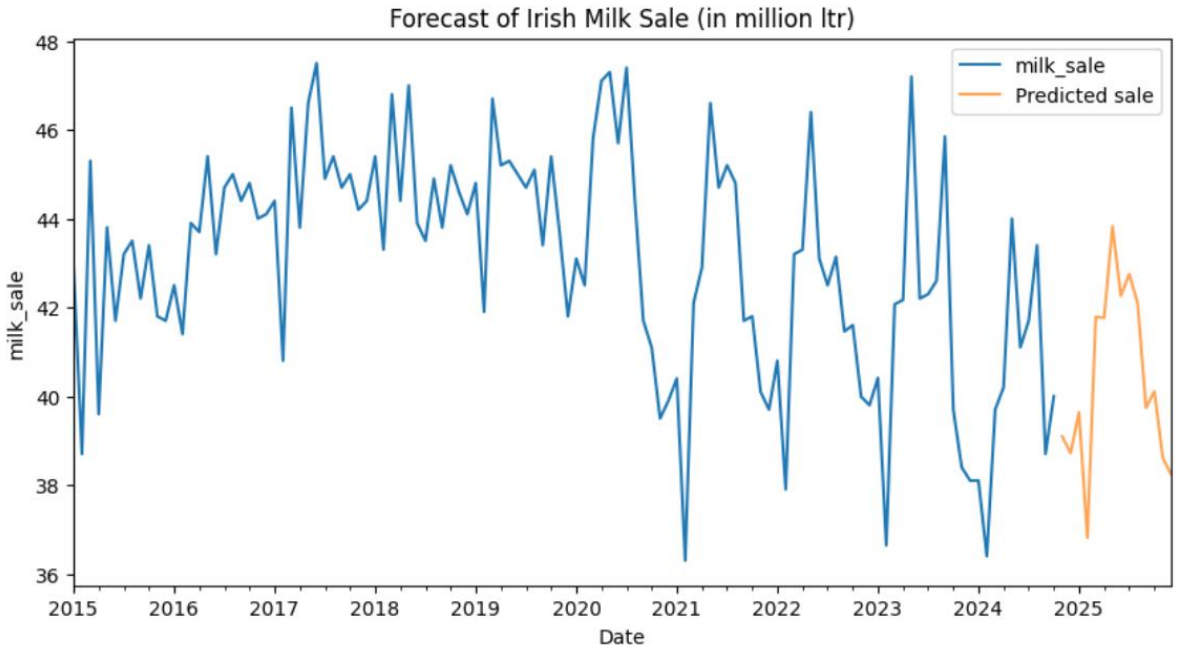
Constant Follow-up
Consistently took the initiative to monitor project progress with other team members, a crucial element in achieving any goal.

At an individual level (self-project work):

1) Searching for the publicly available retail dataset
2) Studying the blogs, research papers and YouTube videos to better understand the concepts particularly Autoregression, Moving Average, ACF, PACF, Stationarity, Decomposition plots, lag operators, ARIMA, SARIMA, SARIMAX models, interpretation of the model parameters etc.
3) Implementing the techniques the chosen dataset.

4. Progress

At an individual level, I was able to complete the forecasting of the milk sales using the SARIMAX model.
In the following image, I have shown that I was able to forecast the milk sales from 2024/11 to 2015/12 (For nearly 1 year period)



At a group level, all other team members were able to execute various time-series models and forecast the sales targets.

5. Challenges and Solutions

At an individual project level, I faced following challenges:

1) The publicly available Irish Monthly Milk Sales dataset contained records from very old years

2) The Milk has several types apart from the “All types”

Because of very old data and absence of any external variable (such as temperature, economic indicators, other factors), the dataset did not have sufficient predictive power. Therefore, to work around this problem, I filtered the dataset to contain records from the year 2015 and limited to milk types to “All Milk”. By making these changes, I was able to get good results.

6.Future Scope

I used the SARIMAX model however there is NO exogeneous variable because the dataset does not have any such. In future, if we get any dataset containing the milk sales data with other external variables then we can use it in our SARIMAX model and the results would be much more accurate and reliable.

7. Learnings

As specified in the TABA question paper, we were able to meet following objectives in relation to our group project.

LO1 -> Analyse, compare, contrast and critically evaluate the characteristics of retail datasets utilised for AI solutions in “Time Series implementation”.
LO2-> Critically assess the challenges associated with implementing Time-Series datasets
LO3-> Critically assess methods and practices for Time-Series Model development
LO4 ->Evaluate, design and implement Forecasting solutions by using various time-series models such as ARIMA, SARIMA, PROPHET etc.

8. Conclusion

As a group project, on a technical front, we were able to learn Time-Series domain in detail and explored various models for forecasting. On the team or project management front, we were able to learn how to plan the deliverables, how to coordinate within the team to achieve the common goal and more importantly the time management is so critical to success of any project.