Executive Report

Impact Project
KPMG
ACME Textiles

 Lucas
 Vignesh
 Rita
 Daniel
 Laufey

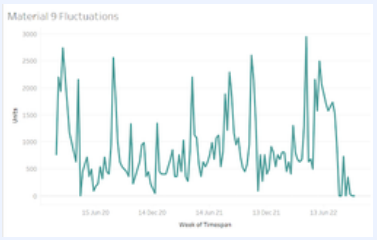
2023

Introduction

This capstone project aims to address the demand forecasting inefficiency in ACME TEXTILE, a company that produces and sells clothing for service professionals. The existing forecasting method, based on subjective assessments, has led to financial inefficiencies. The project aims to develop an advanced forecasting model using historical sales data to improve accuracy and enable optimized production planning and inventory management. The Data Scientist's role is to design and implement this model, surpassing the current accuracy rate of 80% and ensuring long-term financial stability for the company. Code and charts can be found at: https://github.com/notvikke/KPMG_TimeSeries_Forcasting_Capstone

Data

ACME textiles should forecast future values and prepare their production and inventory accordingly. The time series dataset consists of three columns: "year-week", "ID_material," and "Units." The "year-week" column represents the date range from April 2018 to October 2022 in the format YYYY/WW. The "ID_material" column contains the numerical ID of 14 different products, which have been assigned numerical labels for easier analysis. The "Units" column represents the quantity sold in a given week for each material, ranging from 0 to 8316. By analyzing this data, one can identify patterns and dependencies to improve their forecasting and make informed decisions about production and inventory management.



An example of a Distribution Material 9

ACME TEXTILE CORP can improve their demand forecasting by including additional features that influence demand, such as promotions, prices, seasonality, holidays, and economic indicators. These factors can be incorporated into the forecasting model to capture more nuanced patterns and improve prediction accuracy. By considering variables like the number of promotions, consumer price index, and measuring demand by client, the company can better understand the impact on demand and make more precise forecasts. These enhancements will overcome the limitation of weekly sales data and enable ACME TEXTILE CORP to optimize production and inventory management based on more accurate predictions.

Trial and Error

01	STATISTICAL MODELS (ARIMA, EXPONENTIAL SMOOTHING)	• Performed Statistical tests:- ACF, PACF, box-test, and normality tests for data decomposition, checking for seasonality and trends	Accuracy 92%	• Overall the best model on average, recommended for scalability point of view
02	LAZY STRATEGIES (MEAN, MEDIAN)	• A simple dummy model by just predicting the mean/median for each product	Accuracy 77%	• For certain products where the test set (last two months) is drastically different, this simple model outperforms the rest
03	ML MODELS (XGBOOST)	• Implemented XGBoost after trial with carefully feature engineered columns and hyperparameter tuning	Accuracy 83%	• Most ML models tend to overfit to train and gives inaccurate predictions on test, this is mostly due to the randomness in the test data
04	DEEP LEARNING MODELS (RNN)	• RNN being able to use the feedback mechanism from the previous steps enables short and long term dependencies	Accuracy 86%	• RNN is noticeably better than our standard XGBoost model
05	OTHERS PYCARET,PROPHET	• Trial of different approaches which proved to not lead to better results	Accuracy --%	

Overall Accuracy on unseen data using the mean for white noise series: 83%

Models & Evaluation

The objective was to achieve an accuracy higher than 80% with sophisticated forecasting models. The different models tried were statistical approaches like ARIMA (Autoregressive Integrated Moving Average) as well as more advanced deep learning / machine learning techniques. The model which achieves the highest accuracy will be selected, evaluated and deployed. It was decided to evaluate accuracy with MAPE as this is regression problem.

The prediction was to be a 2 point ahead forecast, but whether to calculate the MAPE using 2 or 1 prediction, meaning 1 prediction for 1+2 or pred_1 and pred_2 reamined a question. It was decided to go with the first approach as it would make sense production-wise, if the company wants to forecast the next 2 months from today it is probably because their production has to be arranged bi-monthly. For comparison purposes we will display the (1-MAPE) metric, where 100% means a perfect model. Keep in mind that different metrics and validations techniques have been used for each model, but with the purpose of comparing them and selecting the best ones in our final conclusions we will only look at the test set of the last 8 weeks, this will later be explained again.

Results

After running all models, it became clear that a combination of different models for each product would give the highest accuracy. Regarding feasibility of a deployment, a pipeline could be built with separate models per product. However, this might be difficult to scale. If that is the case, using an ARIMA model for each product should be best for average overall predictions and accuracy.

The accuracy achieved by combining different models that had the best performance on each material respectively was 89.44%. However, creating a different pipeline for each material is not economically feasible recommended for a company the size of ACME. The next best option is choosing the model that on average preformed the best which was an ARIMA which achieved a score of 87%

89.44%

Accuracy

80% › 87%

Recommendations

ACME TEXTILE CORP can improve their forecasting results using weekly sales data by implementing several strategies. Firstly, they can aggregate the weekly sales data into monthly totals, which provides a higher-level perspective and eliminates noise and fluctuations associated with weekly data. Monthly data offers a smoother and more stable basis for forecasting models. Additionally, incorporating additional features that influence demand, such as promotional activities, product price, seasonality, holidays, or economic indicators, can enhance accuracy. By including these factors as input features in the forecasting model, the company can capture more nuanced patterns and improve predictions.

Implementing these strategies allows ACME TEXTILE CORP to overcome the limitation of having only weekly sales data, optimizing production and stock management accordingly. Furthermore, deploying ARIMA and exponential smoothing models on a cloud based MLOps platform can automate the entire lifecycle of the forecasting models. This approach enables the business to generate accurate and up-to-date sales forecasts, make informed decisions, and optimize operations effectively. By leveraging these strategies and technologies, ACME TEXTILE CORP can enhance their forecasting capabilities and achieve better outcomes in their production and stock management processes.

Business Case

Using the ARIMA model, the analysis reveals a decrease of 15,500 wrongly allocated units, comprising 13,000 overproduced and 2,500 underproduced units. This reduction in misallocation results in a decrease of variable and fixed costs by €3,584,706, as well as avoiding an opportunity cost of €175,804. In Scenario 1, by reducing overproduction, ACME Textile optimizes its production process, avoids excessive units, and saves operational costs.

Wrongly Allocated Units

44000 › 28500

Over-produced units decrease: 13000
Under-produced units decrease: 2500

In Scenario 2, ACME Textile addresses underproduction, which occurs when the actual demand is not met. By implementing the ARIMA model, ACME Textile identifies an underproduction of 2,500 units, resulting in missed sales opportunities. By decreasing underproduction, ACME Textile captures additional sales revenue (€897,500) and profit (€179,500), improving market competitiveness and customer satisfaction.

The implementation of the advanced forecasting model not only enhances sales performance but also positively impacts financial results. Accurate demand forecasting ensures optimal production levels, meets customer demands, and drives business growth for ACME Textile.