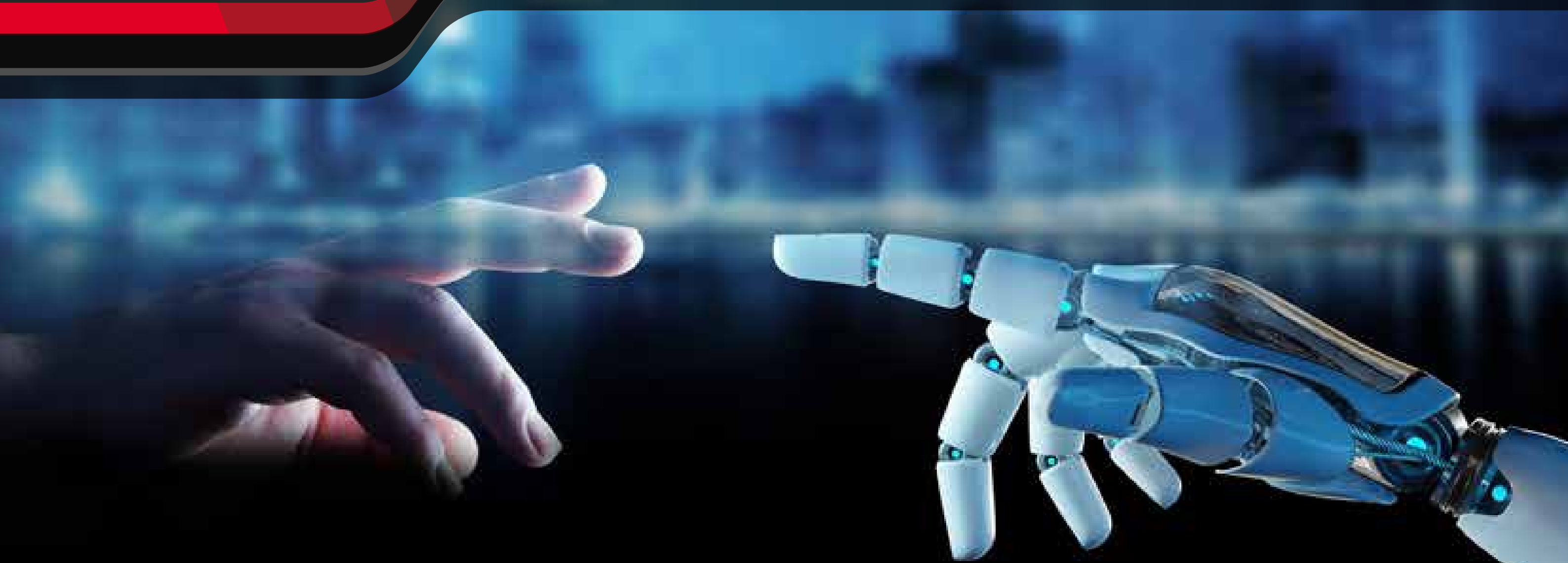
Chartered Institute of Global WFM, USA & Canada

GWFM® GLOBAL WORKFORCE MANAGEMENT

THE BIGGEST WFM NETWORK IN THE WORLD

Global Chapters : APAC | ANZ | EMEA | NA | LATAM



WHITE PAPER

MACHINE LEARNING BASED FORECASTING ENGINE

INDEX

GWFM

Executive Summary

The quest is to find the nearest value to the actual target value. An analogy will be like shooting an arrow to a target inside a dark room, we never know if it will hit the bull's eye or not, but we strive to hit as near as possible. After we shoot the arrow the light inside the room is switched on and then we come to know how close we are to the bull's eye. Then the light is switched off again. Next time when we are going to shoot an arrow, remembering how close we were, we do adjustments to hit the bull's eye. In a similar context, most companies and client have been looking for a forecast method or model that can deliver accurate forecasting, however after evaluating the forecasting model that exists in the market I find limitation which has pushed me to think out of the box and check what can be done to help the existing or future clients.

Some forecasts are built on past data values of a variable. In such cases, the past patterns repeat, and we get the advantage of using the past trends of the data. In many cases we don't have past values, so we use some time-tested methods like averaging to arrive at a possible forecast value. In some cases, we do use a cocktail of both past and most recent trends to arrive at a forecast value.

Some examples of method selection:

- Data where past values do show repeatable trends, we do use methods like **"ARIMA, Seasonality with trend methods, Holt winters (LTS)".**
- Data where past trends are not very influential to the future data, we use averaging methods like **"Moving Average, Simple Exponential smoothing".**
- Data where we want to use past trends as well as latest current trends, we use methods like **"LSTM, ANN".**

Every forecasting method has its pros and cons, the best method is to test several different methods and select the one which gives values nearest to the target values and it is a continuous process. In this particular case, we saw value in implementing LSTM as we wanted to use past trends as well as the latest current trends to derive the forecast value and we saw the merit when we implemented the model.

Project Description

Build forecasting engine using latest technology and scientific approach, that can deliver a demand accuracy between +/- 5% and required minimum manual intervention.

Situation

The existing process in companies has limited forecasting and capacity planning capabilities. Lack of scientific approach for Demand and Supply planning, the planning is done basis workload calculation and not on any methodical approach, which lead to capacity gaps of 15 to 20% on few days, staffing gap, revenue impact, customer satisfaction and employee satisfaction issues. The process make limited consideration of the offered volumes, future projections, hiring impact and employee dissatisfaction. The service delivery team normally call employees in standard shifts and in case of inadequate work volume, the employees are asked to send back or close the shift early with the expectation of covering the lost hours on other days.

Apart from this there were other challenges that I have seen like:

- Data had been processed manually also effecting quality and volume of data
- Data constraints: Companies have data only for 18 months, limited insight on volume variance from Ops team, only monthly data available for support from other teams and shift call off details unavailable
- Outputs of the forecasting process are manually copied into other tools as inputs for requirement calculation and further reporting

Task

The task was to identify, build and deploy forecast model that can help the service delivery team with resource planning (including workload calculations), address issues related to planning, having optimal resource staffing and ensure volumes are addressed within Service Level. Factors to be considered were long term outlook, arrival pattern and seasonality.

Forecasting is vital first step in workflow cycle, affecting most of the processes coming after it: capacity planning, resource management and scheduling. Success of each of those directly relies on the quality of the forecast, because forecasts are used to calculate staffing which Conduent deploys to handle incoming transactions/tickets/volume which in turn generate revenue for the company.

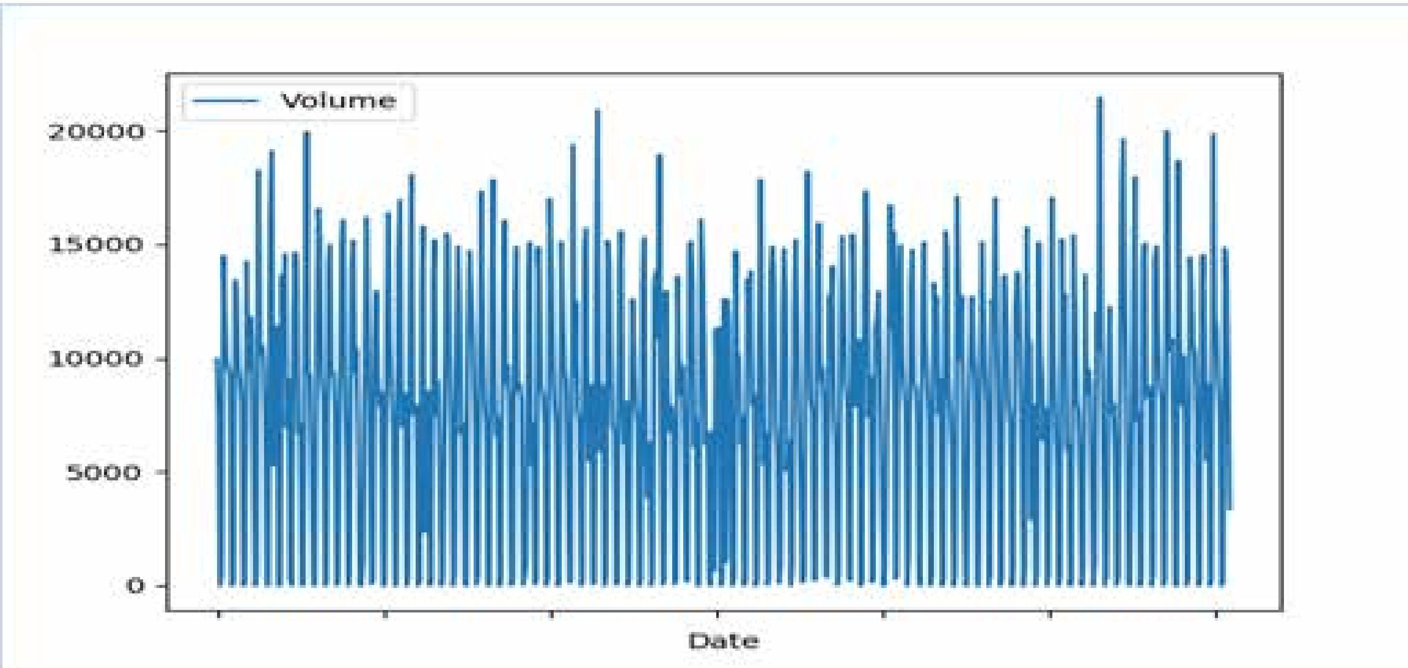
Activity

Analyze

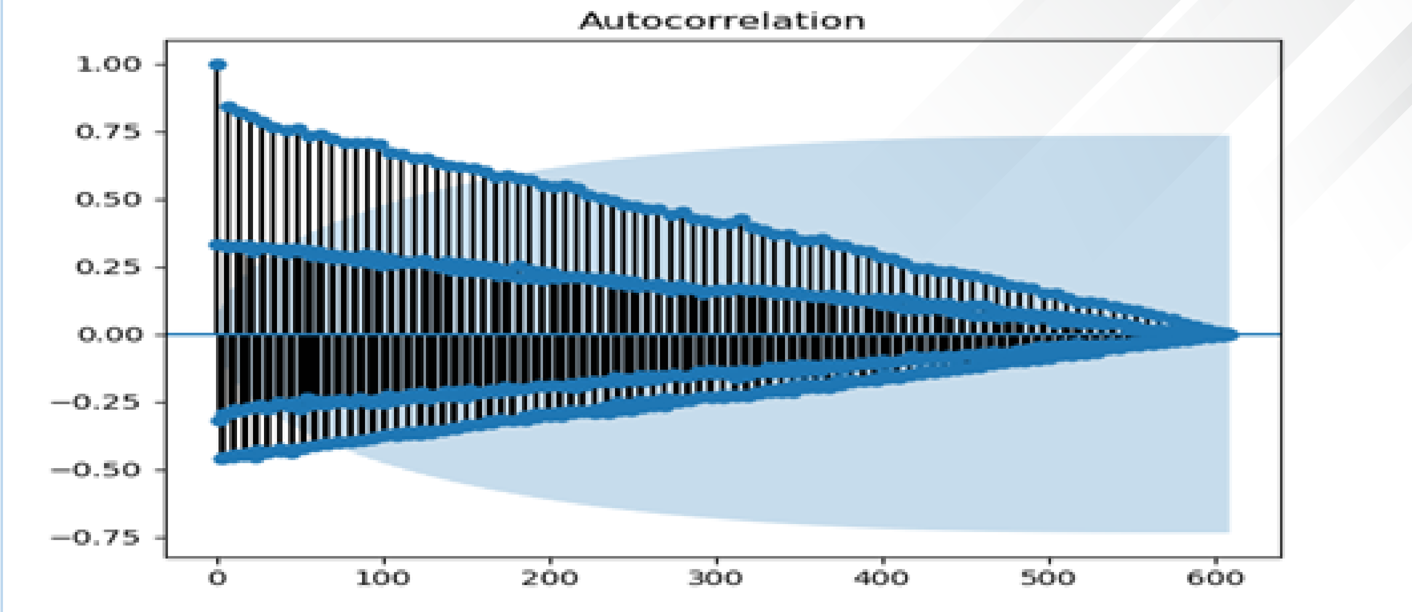
The approach was to first understand the workflow, collocate and analyze the data to understand business as usual volume, seasonality and growth impact. Since in this case there was very less data available to analyze the trend took statistical approach to run data test using various analytics tools and the observations derived from these tests were contrary to being right and indicated the following:

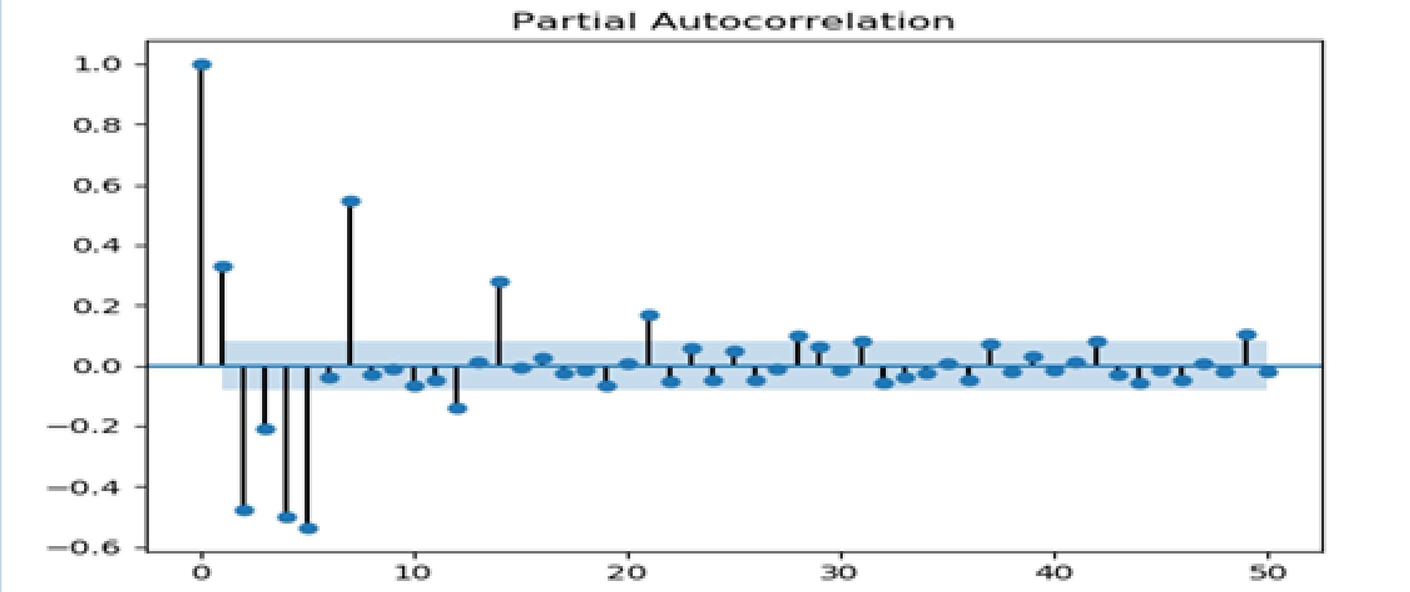
- Periodicity pattern plot showed irregularity in the patterns (at someplace it was low and at some places it was high)

Periodicity pattern plot showed irregularity in the patterns (at someplace it was low and at some places it was high)

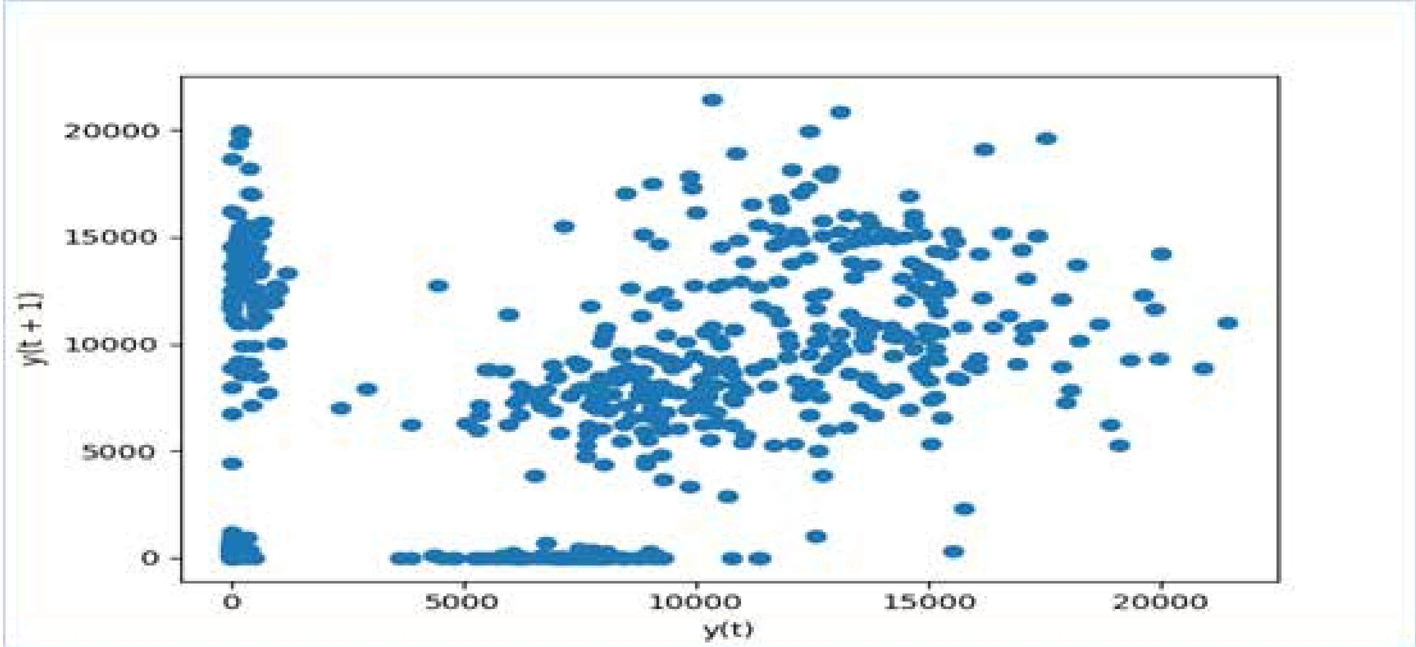


The auto-correlation and partial auto-correlation indicated noisy patterns due to the effect of some promotional /surprise events. Even that seemed to have some level of non-linearity

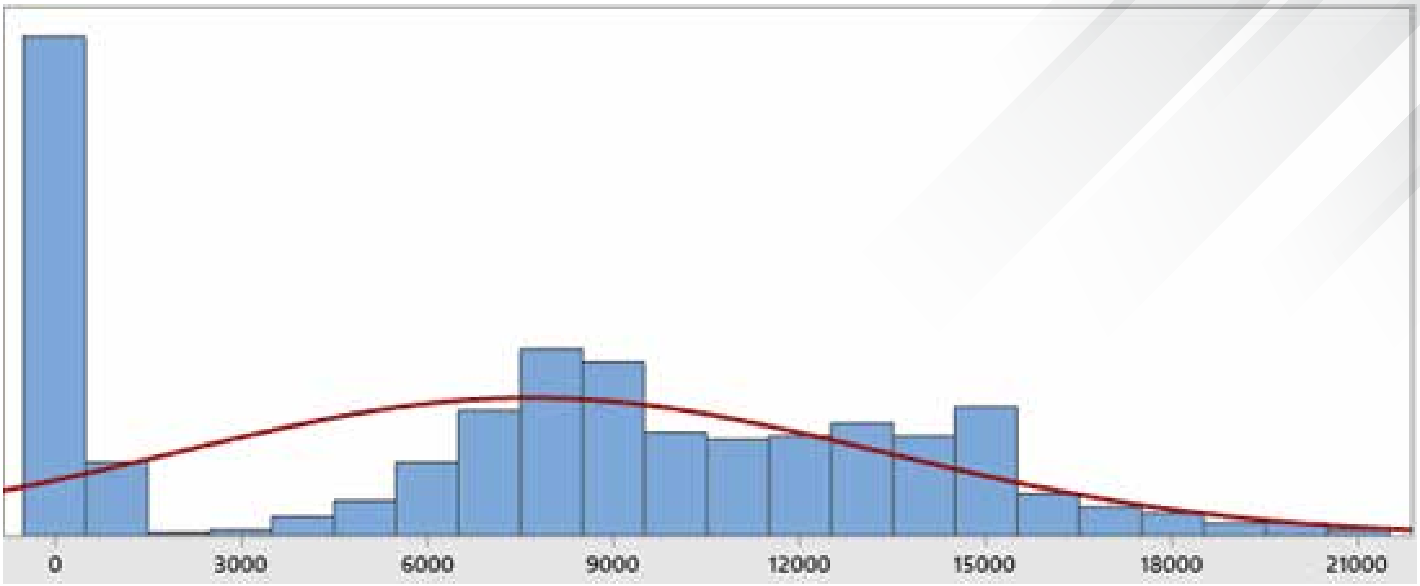




Lag-Scatter plot also indicated some level of non-linearity in the data. It also reflected the less amount of periodicity (as a lot of sparse data points were there in the plot-diagram)



Standard deviation with complete data was ~5700 & the P value was less than 0.05 with limited insight on data seasonality and spike in volumes



Anderson-Darling Normality Test	
A-Squared	16.7
P-Value	<0.005
Mean	7566
StDev	5700
Variance	32175200.3
Skewness	-0.02276
Kurtosis	-1.11121
N	600
Minimum	0.0
1st Quartile	483.5
Median	7991.0
3rd Quartile	11986.6
Maximum	20990.0

As a next step, the data was normalized using exponential smoothing and eliminating few data points where variance was more than 150%

The team also did a parallel research to understand various forecasting models that are available in the market, shortlisting the best used case to understand that model that will fit the need of the client in the current environment and help us with automation. Below is the comparison of various models that are currently available:

Current Excel based Models	Traditional ML models (IEX/Bluepumpkin)	Deep Learning/Neural Network (Proposed)
Short-term (weekly, daily forecasts) with complicated macros	ARIMA performs well for short-term forecasts, but for long-term forecasting and for large datasets , variance in the predicted output increases	Deep Learning based Systems are effective in long term data
Manual adjustments to forecasts to account for seasonality	Traditional ML models like ARIMA do not support non-linearly co-related data sets. Data coming from multiple sources generally results in non-linearity	Deep Learning based System Supports Non-Linearity
Forecasts are not optimal due to manual errors		Other proposed approaches include Multi-level machine learning models (prediction from previous levels as inputs to next level model to improve accuracies)
Manual collection of data for input to the excel		
Large excel files due to historical data storage in files		

WFM Forecasting – Why Deep Learning based System?

All – available WFM systems (i.e., IEX/Blue Pumpkin, etc.) use ARIMA model for forecasting. Facts from Research Papers (ARIMA Vs Deep Learning):

WFM Forecasting – Why Deep Learning based System?

1. Deep Learning based Systems are effective in long term data, i.e. when the test dataset is large.

1. ARIMA for instance gives more importance to immediate data points in the test set and tries to perform well for them but as we get far we see a larger variance in the predicted output

2. Deep Learning based System Supports Non-Linearity:
1. Due to the dynamic nature of the time series data, we often see that there is a non-linear co-relation between the past and the current data points.
2. Data coming from multiple sources generally results in non-linearity.
3. Moving average based techniques do not support non-linearity

Design & Implement Solution:

After analyzing the data and looking at the market research information the project was divided into two phases:

Phase - 1	Phase - 2
Building manual prototype to test the best fit model with existing resources and test the result	Fully automate the model that had capabilities to adjust seasonality factor manually at the last minute incase needed

Phase 1

The demand planning model that were used - ARIMA, Exponential smoothing, Multi-regression, Holt winter, Moving average and weightage average.

Approach 1

(Week 41 to Week1): Applied the data on the above forecasting models using internal ources and used external vendors to understand which model works the best.

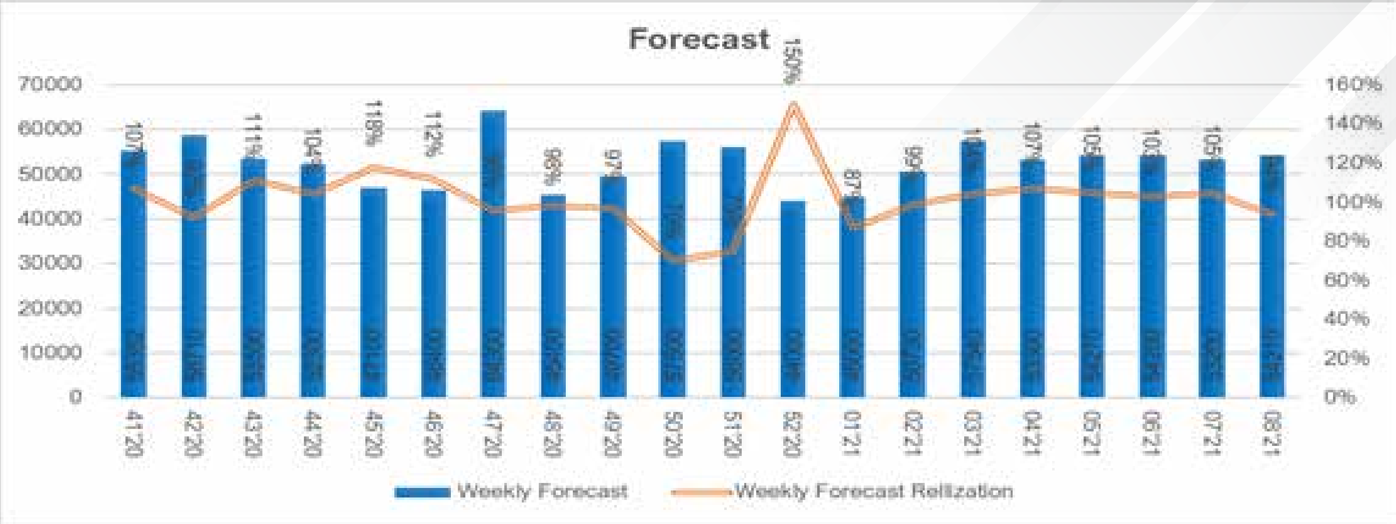
Monitor Results and Process: The first test was passed with combination of two different models, with forecasting variance of 15%+/- however we needed to get to 10%+/- or below to show better cost goodness hence the team worked on further enhanced model before we invest time and cost on automation.

Approach 2

(Week 2 to Week8): Totally different approach of combining three different models, the test was done using "Minitab" tool and two different forecast models were combined using Exponential smoothing + multiple regression + decomposition model.

Monitor Results and Process: The test passed with combination of three different models with a forecasting variance of 5 to 7% +/-. However, the process was quite tedious as we first removed data anomalies in excel using exponential and then fed the data into "Minitab" to run multiple regression model. The results were again downloaded fed back into Minitab to run Decomposition model to get the final output.

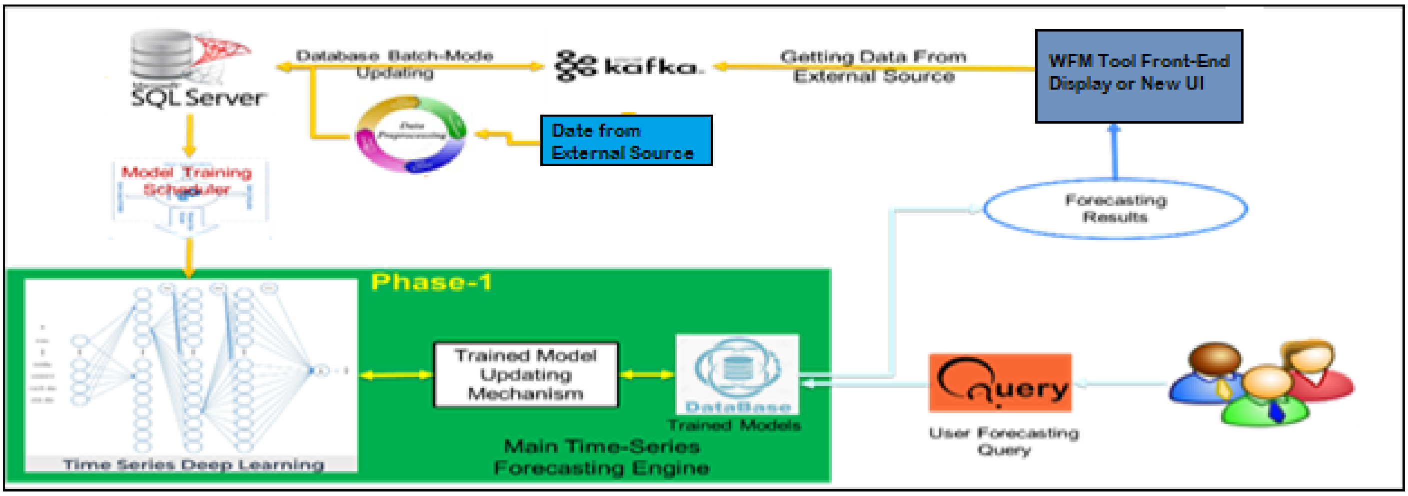
Forecasting Results:



Phase - 2

Standardize and Share Success: Post deploying the above models the same case study was discussed with research team to automate the complete forecasting engine to avoid manual intervention and the lab team used LSTM (Long Short-Term Memory) forecasting model to automate the complete process. Python was used as a code language for the programming of this model and automation BOTS were used to feed the data into the forecast engine.

The reason for using LSTM was that it is a neural network based model and uses deep learning capabilities. Deep Learning enabled the BOT to continuously evolve the model, basis the change in the pattern, without any manual intervention.



Results:

Reduced forecast variance to 7% +/- that has improved operational efficiency and reduction in $0.5 Million/100 FTE. The automation of tool has helped operation take informed operational decision and improve effectiveness. It has also helped Workforce Management team to reduce manual effort, improved data and helped deliver better forecast accuracy. Considering that this pilot has yielded results beyond expectations, it is being replicated across multiple accounts as a best practice.

About GWFM Chartered Body Canada & USA

Global Workforce Management Forum is the 1st WFM Forum in the world initiated. GWFM form is a nonprofit professional organization and Professional Global Charted Body, comprising of a group of Senior WFM & HR professionals committed to promoting the WFM movement in the global platform and enhancing the capability of workforce management & human resource professionals to compete globally and thereby creating value for society. GWFM is committed to the development of the workforce through education, training, research, and experience sharing. GWFM is one of the fastest-growing professional forums in the world and a trendsetter where we discuss current and future happenings in the global business.

Established in 2014 June, GWFM is an autonomous, not-for-profit professionally managed organization, playing an Evangelist role in grooming Leaders for Future. GWFM has 20,000 members representing, Top Multinationals, Public & Private organizations including Top Universities spread across the world. GWFM has had an exciting journey of over two and a half years. GWFM has been stupendously enriched by the learning from top industry leaders.

The Global WFM (GWFM) is the Professional Body for WFM and HR professionals across the world. GWFM is a not-for-profit organization that exists to promote greater awareness, understanding, and innovation in Workforce Management is fully supported in its mission by leading organizations in the sector. The GWFM provides members with the opportunity to network with other professionals in the industry, to learn about career development opportunities, to share knowledge and expertise, to access Continuous Professional Development programs, and to carry the official designation of membership.

Reachout for WFM skilling and up-skilling platform: www.gwfmlearning.online

www.globalwfm.com

Author

Arvind Shukla,

Head of Workforce Workforce Service Delivery,
Program Management & Sales Operations
APAC & North Americas, Conduent

GWFM