Demand Forecasting For Industrial Products

Presentation by:
Aditi Kannan
SSN College of Engineering

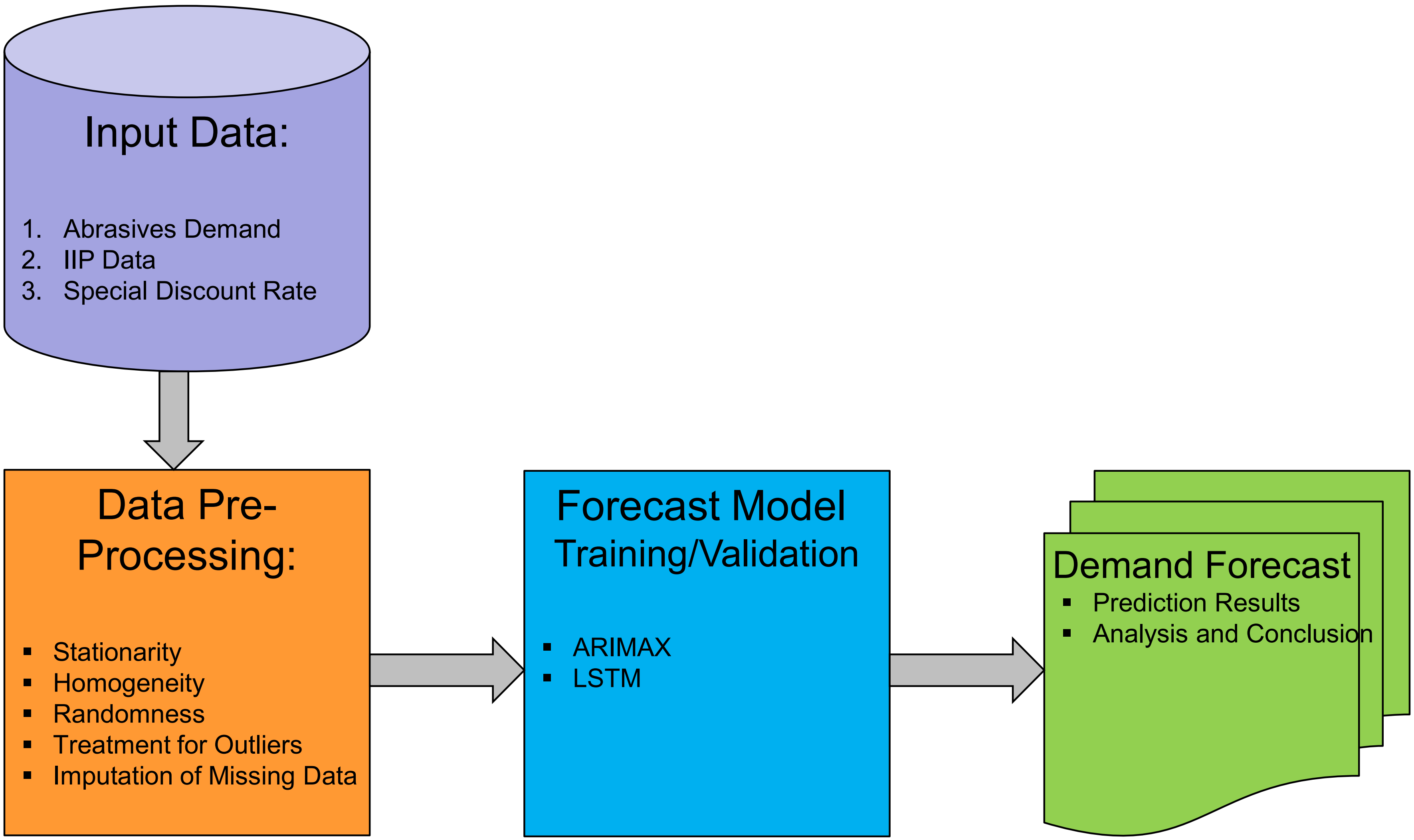
Objective/Motivation

- Any organization/industry faces several internal and external risks. The adverse effects of risks can be lessened by determining the demand or sales prospects for its products and services in future.
- “Demand estimation (forecasting) may be defined as a process of finding values for demand in future time periods.”

Literature Review

Author	Title of the paper	Conference/ Journal Publication	Description
Bontempi, G., Taieb, S. and Borgne, Y.	Machine Learning Strategies for Time Series Forecasting	Lecture Notes in Informatics (LNI) - Proceedings Series of the Gesellschaft für Informatik (GI) Volume P-295 ISBN 978-3-88579-689-3 ISSN 1617-5468	Review of machine learning techniques in time series forecasting
Mehrmolaei, S. and Keyvanpour, M. R.	Time series forecasting using improved ARIMA	2016 Artificial Intelligence and Robotics (IRANOPEN), Qazvin, Iran	The authors propose a novel approach to improve ARIMA model by applying a mean of estimation error for time series forecasting. Experimental results indicate that the proposed approach can improve performance in the process of time series data.
Siami-Namini, S., Tavakoli, N. and Siami Namin, A	A Comparison of ARIMA and LSTM in Forecasting Time Series	2018 17th IEEE International Conference on Machine Learning and Applications (ICMLA), Orlando, FL, USA	The authors investigate the superiority of deep learning-based algorithms for forecasting time series data, such as "Long Short-Term Memory (LSTM)", to the traditional algorithms.
Wanchoo, K.	Retail Demand Forecasting: a Comparison between Deep Neural Network and Gradient Boosting Method for Univariate Time Series	2019 IEEE 5th International Conference for Convergence in Technology (I2CT), Bombay, India	The author implemented two machine learning techniques, Deep Neural Network (DNN) and Gradient Boosting Method (GBM) for univariate time series sales data at store-day level of a German retail giant
Bandara, K., Shi, P., Bergmeir, C., Hewamalage, H., Tran, Q. and Seaman, B.	Sales demand forecast in e-commerce using a long short-term memory neural network methodology	Neural Information Processing: 26th International Conference, ICONIP 2019, Sydney, NSW, Australia	The authors attempt a model which conditions the forecast of an individual time series on past behaviour of similar, related time series. This is achieved by globally training a Long Short-Term Memory network (LSTM) that exploits the nonlinear demand relationships available in an E-commerce product assortment hierarchy

Forecasting System



ARIMAX Model

Linear regression equation for ARIMAX :

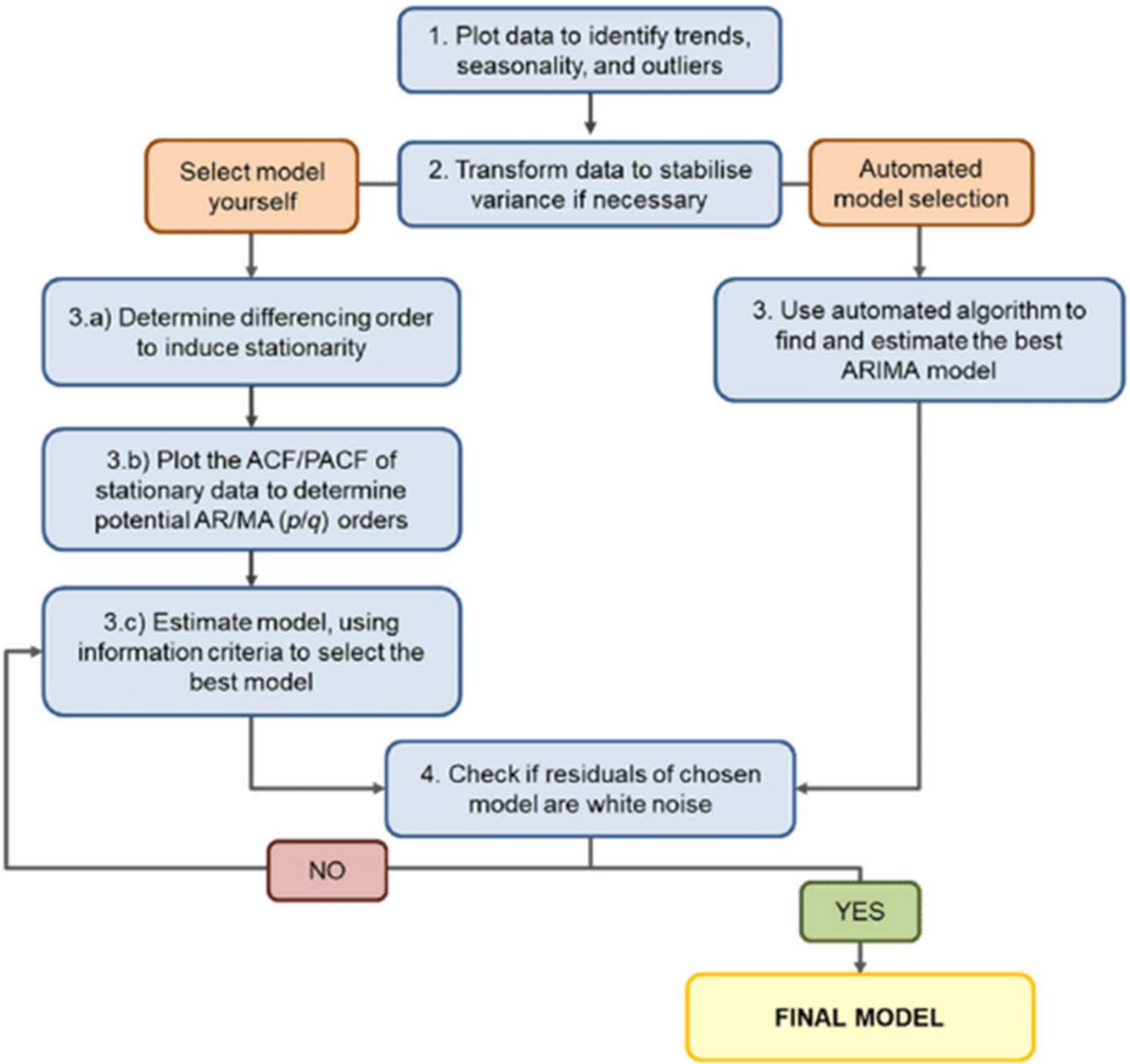
$$Y_t = a_o + a_1 Y_{t-1} + \dots + a_p Y_{t-p} +$$
$$\theta_o + \theta_1 \varepsilon_{t-1} + \dots + \theta_q \varepsilon_{t-q} +$$
$$\beta_o + \beta_1 X_1 + \dots + \beta_k X_k$$

- Yt is the dependent variable at time t.
- a1 to ap are the autoregressive coefficients of order p.
- θ1 to θq are the moving average coefficients of order q.
- ε_{t-1} to ε_{t-q} are the first q moving average terms
- β represents the differencing coefficients of own time series

ARIMAX Model

- The autoregressive part (AR) forecasts the variable of interest by using the past values of the variable
- The integrated part (I) represents the differencing of raw observations to allow the time series to become stationary.
- The moving average part (MA) uses past forecast errors in a regression-like model. The past forecast error is the difference between the actual data and fitted values
- X represents exogenous variables that are independent of time-series but have effect on the time-series data. Exogenous variables are often used to capture the effect of external factors that may influence the behavior of the dependent variable

ARIMA Model Selection



Let $B(x_t)=x_{t-1}$ denote the lag operator.

Let us define $z_t=y_t-y_{t-1}$ as the response.

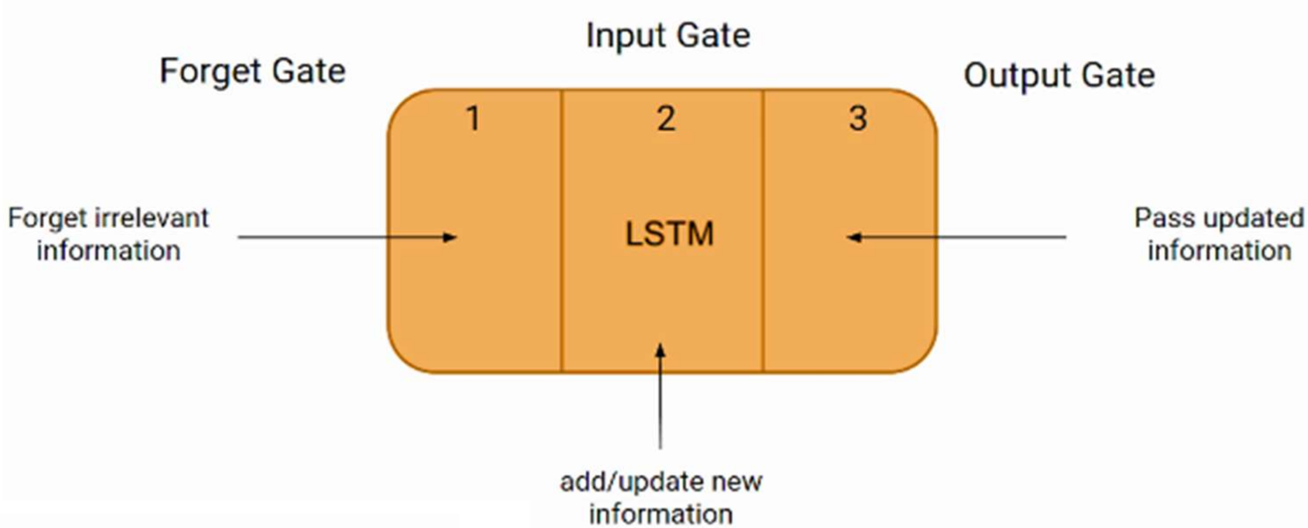
The ARIMA (2,1,1) can be defined as

$$z_t = \phi_1 B(z_t) + \phi_2 B^2(z_t) + \epsilon_t - \theta_1 B(e_t)$$

LSTM Model

- LSTMs work by maintaining a cell state that can store information over time, selectively updating and resetting the cell state based on input data. The cell state is controlled by gates, which are a set of sigmoid neural network layers that decide how much information to keep, how much to forget, and how much to output at each time step.
- LSTM architecture consists of four main components: the input gate, the forget gate, the output gate, and the cell state
- Input gate: Determines how much of the input should be added to the cell state.
- Forget gate: Determines how much of the previous cell state should be forgotten.
- Output gate: Determines how much of the cell state should be output.
- Cell state: The memory of the LSTM that stores information over time.

LSTM Model



Output Gate:

- $o_t = \sigma\,(x_t * U_o + H_{t-1} * W_o\,)$

$H_t = o_t * \tanh(C_t)$

Output = Softmax(H_t)

Forget Gate:

- $f_t = \sigma\,(x_t * U_f + H_{t-1} * W_f\,)$

- Xt: input to the current timestamp.
- Uf: weight associated with the input
- Ht-1: The hidden state of the previous timestamp
- Wf: It is the weight matrix associated with the hidden state

$C_{t-1} * f_t = 0$ ...if $f_t = 0$ (forget everything)

$C_{t-1} * f_t = C_{t-1}$ ...if $f_t = 1$ (forget nothing)

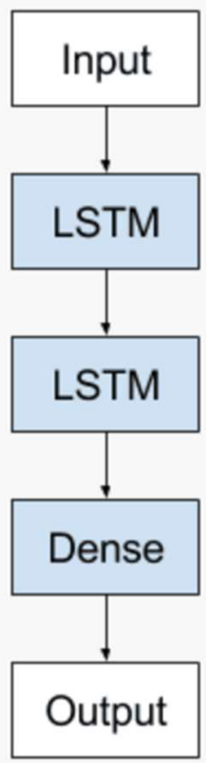
Input Gate:

- $i_t = \sigma\,(x_t * U_i + H_{t-1} * W_i\,)$

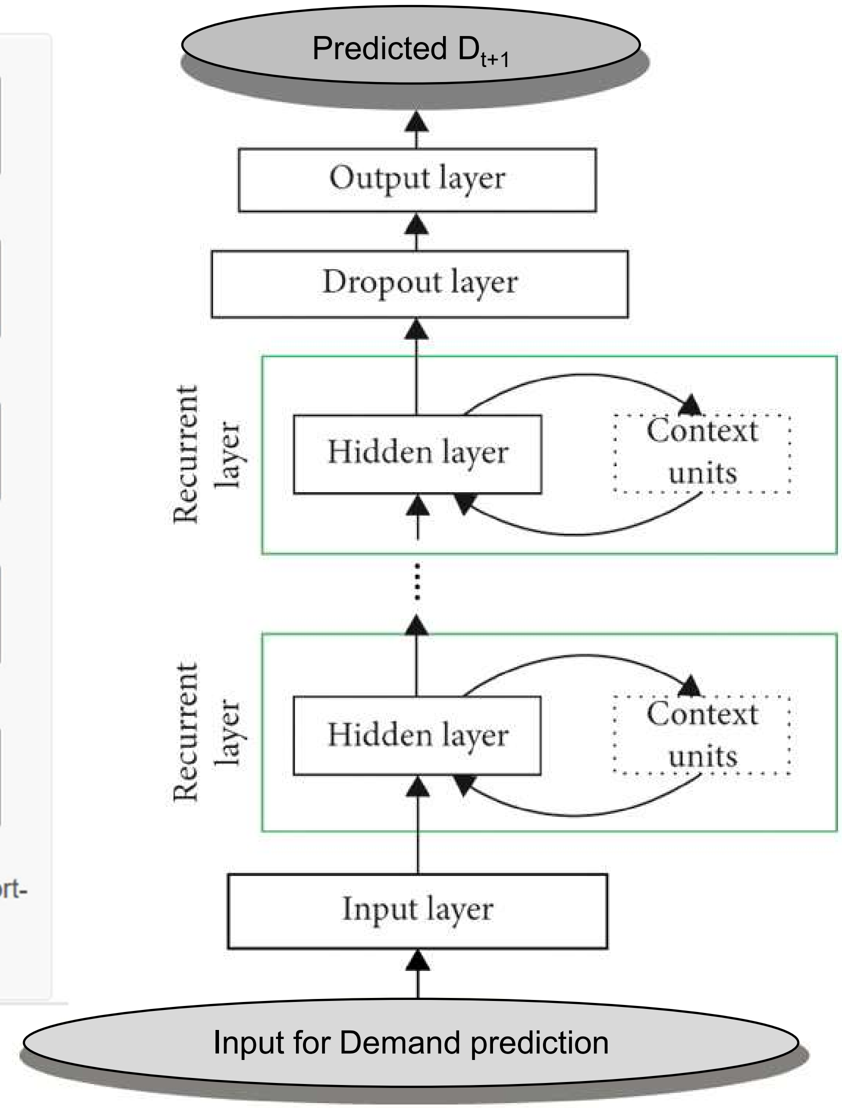
- Xt: Input at the current timestamp t
- Ui: weight matrix of input
- Ht-1: A hidden state at the previous timestamp
- Wi: Weight matrix of input associated with hidden state

- $N_t = \tanh(x_t * U_c + H_{t-1} * W_c\,)$ (new information)

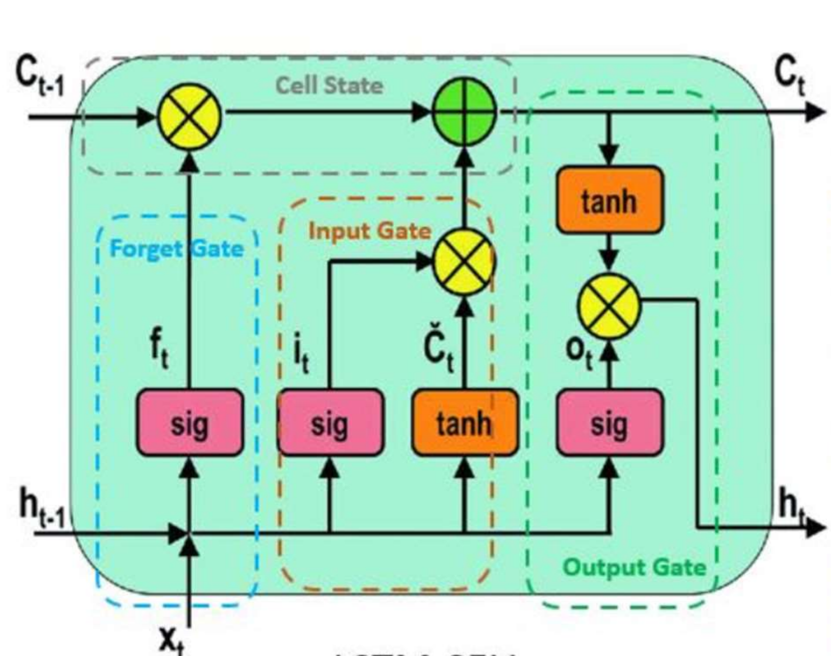
LSTM Model



Stacked Long Short-Term Memory Archicture



Number of Hidden Layers : 50
Dense Layer : 1
Loss Function : MAE
Optimizer : Adam



LSTM CELL

Performance Metrics

- Mean Absolute Percentage Error

Formula

$$M = \frac{1}{n}\sum_{t=1}^{n}\left|\frac{A_t - F_t}{A_t}\right|$$

M = mean absolute percentage error

n = number of times the summation iteration happens

A_t = actual value

F_t = forecast value

MAPE	Interpretation
< 10%	Very Good
10% - 20%	Good
20% - 50%	OK
> 50%	Not Good

Dataset Details

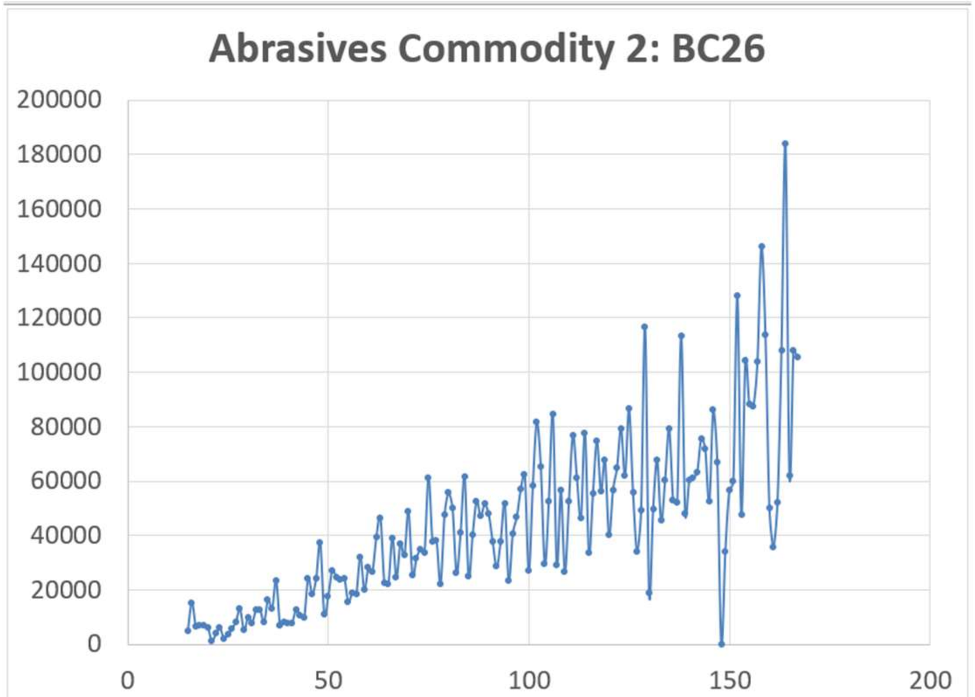
Monthly Time-series data on

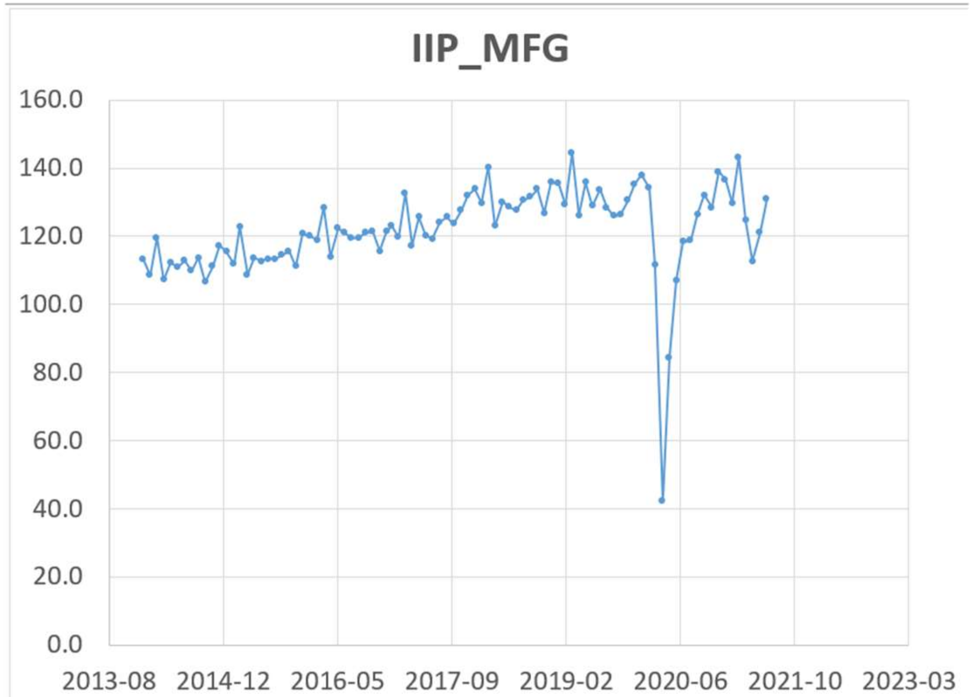
- Abrasives Demand
- Indices of Industrial Production
- Monthly Special Discount Rate

Data Availability Statement

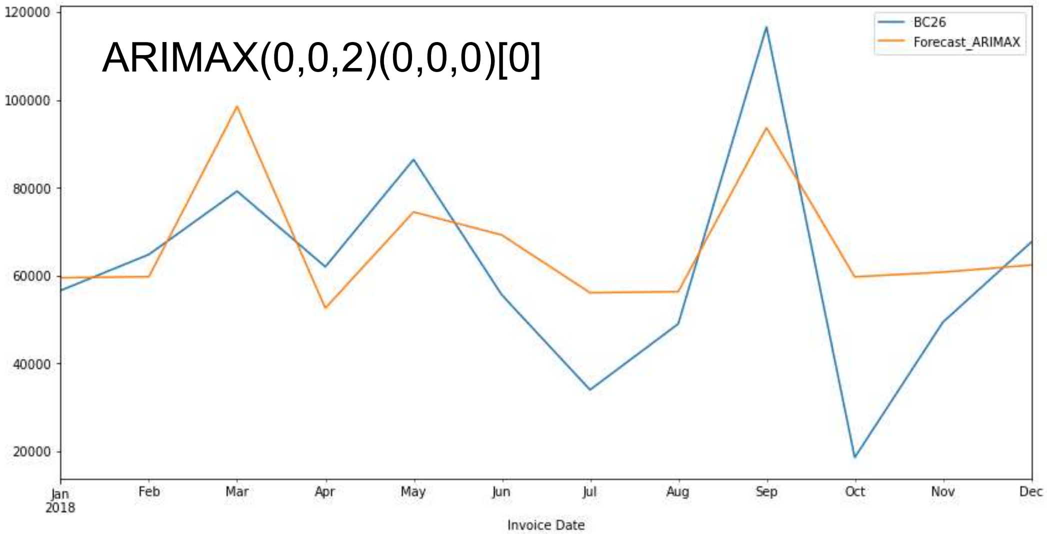
Sl. No	Description	Availablity	
		From	To
I.	Abrasives Demand		
	a) BC26	"2009-03-01"	"2021-11-01"
	b) BC27	"2009-03-01"	"2021-11-01"
II.	Index of Industrial Production (IIP)		
	IIP_MFG	"2014-01-01"	"2021-07-01"
III.	Special Discount Rate		
	a) BC26	"2009-03-01"	"2021-11-01"
	b) BC27	"2009-03-01"	"2021-11-01"

Time-Series Plot of Dataset

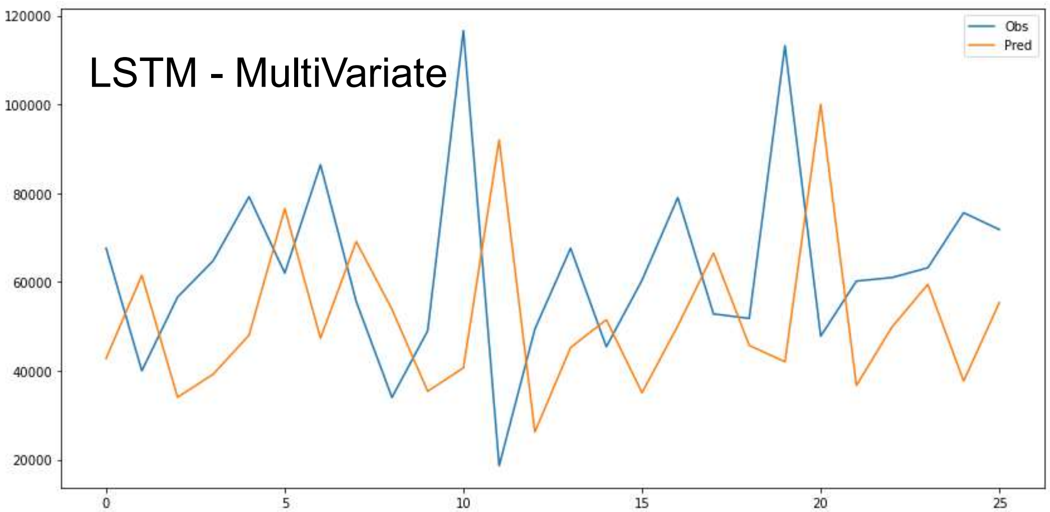


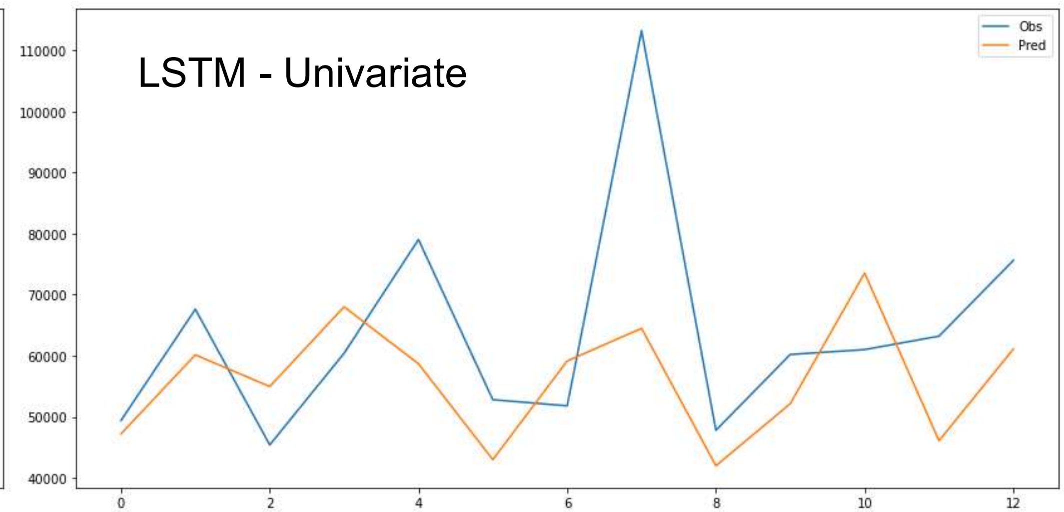


Results



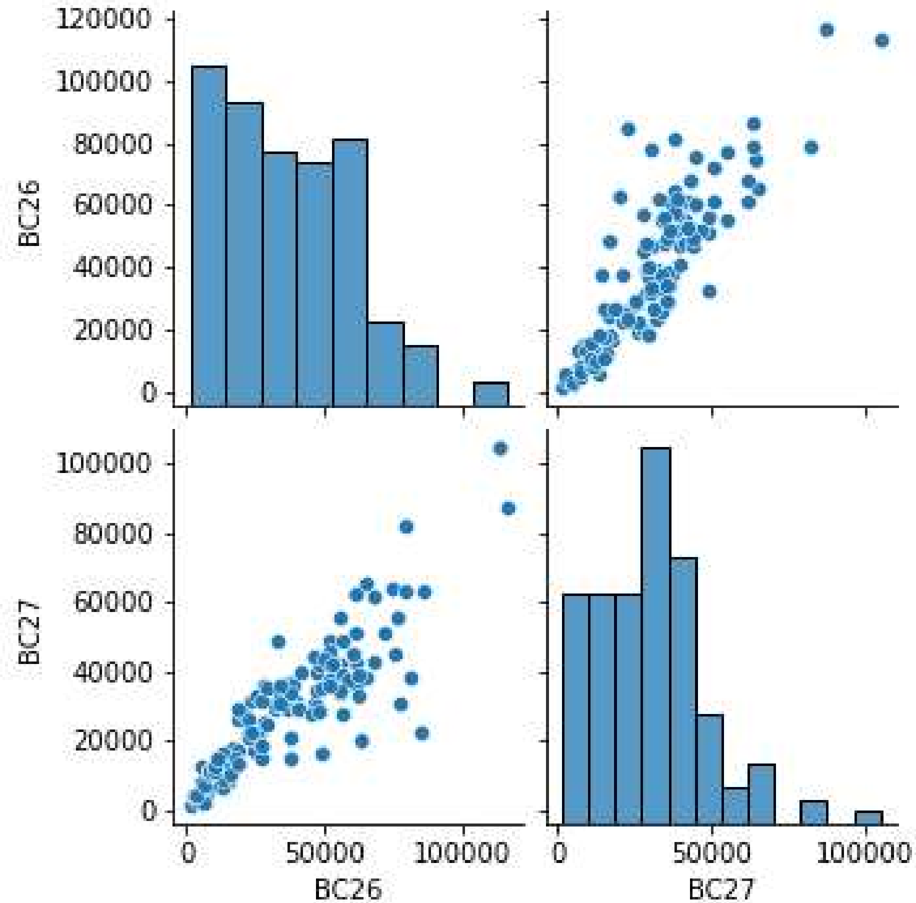
Sl. No.	Model	MAPE
1	ARIMAX	36.852
2	LSTM - Multivariate	100.77
3	LSTM - Univariate	95.91





Conclusion & Work to be done

- ARIMAX model performs better than LSTM models
- The external variables such as IIP and special discount rate does not have any correlation with demand.
- The demand time series is affected by discount provided by the seller and the discount provided does not have any specific annual pattern.
- There is a possibility of finetuning parameters of LSTM models.



References

1. P. Pawar, S. Hatcher, L. Jololian and T. Anthony, "Demand Forecasting using Machine Learning," 2019 SoutheastCon, Huntsville, AL, USA, 2019, pp. 1-8, doi: 10.1109/SoutheastCon42311.2019.9020393.
2. M. E. Hoque, A. Thavaneswaran, S. S. Appadoo, R. K. Thulasiram and B. Banitalebi, "A Novel Dynamic Demand Forecasting Model for Resilient Supply Chains using Machine Learning," 2021 IEEE 45th Annual Computers, Software, and Applications Conference (COMPSAC), Madrid, Spain, 2021, pp. 218-227, doi: 10.1109/COMPSAC51774.2021.00040.
3. K. Wanchoo, "Retail Demand Forecasting: a Comparison between Deep Neural Network and Gradient Boosting Method for Univariate Time Series," 2019 IEEE 5th International Conference for Convergence in Technology (I2CT), Bombay, India, 2019, pp. 1-5, doi: 10.1109/I2CT45611.2019.9033651.
4. Bontempi, G., Taieb, S. and Borgne, Y. (2013). Machine Learning Strategies for Time Series Forecasting. Brussels, Belgium: SpringerVerlag. Available at: https://www.academia.edu/20245534/Machine_Learning_Strategies_for_Time_Series_Forecasting.
5. Bandar, K., Shi, P., Bergmeir, C., Hewamalage, H., Tran, Q. and Seaman, B. (n.d.). San Bruno, USA. Available at: https://www.academia.edu/38165760/Sales_Demand_Forecast_in_Ecommerce_using_a_Long_ShortTerm_Memory_Neural_Network_Methodology.
6. S. Mehrmolaei and M. R. Keyvanpour, "Time series forecasting using improved ARIMA," 2016 Artificial Intelligence and Robotics (IRANOPEN), Qazvin, Iran, 2016, pp. 92-97, doi: 10.1109/RIOS.2016.7529496.
7. S. Siami-Namini, N. Tavakoli and A. Siami Namin, "A Comparison of ARIMA and LSTM in Forecasting Time Series," 2018 17th IEEE International Conference on Machine Learning and Applications (ICMLA), Orlando, FL, USA, 2018, pp. 1394-1401, doi: 10.1109/ICMLA.2018.00227.

Thank you !