# Walmart Sales Forecasting Accomodating Inflation Rate

Laksha Sekar, School of Computer Science and Engineering, Vellore Institute of Technology, Chennai, laksha.s2020@vitstudent.ac.in, Yashika Singh, School of Computer Science and Engineering, Vellore Institute of Technology, Chennai, yashika.singh2020@vitstudent.ac.in Dr Praveen Joe I R, School of Computer Science and Engineering, Vellore Institute of Technology, Chennai, praveen.joe@vit.ac.in

***Abstract*—Sales forecasting, in general, is the process by which a company gets to predict the sales that happen in the future based on the past sales of the products and services of the company, enabling them to predict their revenue. It helps the leaders of the company or organization make important decisions or frame strategies for the organization's future growth. The classic Sales Forecasting algorithms of the current times don't include the important inflation factor while calculating future sales. It is important to note that inflation rates reduce consumer purchasing power, which in turn affects the sales of a company and in turn, the revenue, this consumer purchasing power reduces consumer spending directly. This paper aims to solve that problem by including inflation rates while calculating sales revenue for the future. The Walmart sales dataset is used to forecast sales in this paper. The Dynamic Generalized Linear Model (DGLM) is used in this paper for forecasting sales.**

***Index Terms*—sales, forecasting, inflation, DGLM, Walmart sales dataset.**

## I. Introduction

IN order to make informed judgments about future investments, production, staffing, and overall strategy, companies have to use sales forecasts, making it an integral part of business planning. Sales forecasting helps in overall budgeting for the next financial year for an organization by calculating the revenue generated, the approximate cash flow in and out of the organization and an approximate profit that could be generated. This approximation of revenues by sales forecasting could help the decision makers to implement better strategies to meet up with the demand of certain products in the stores by efficient product production planning, which would in turn require strategic resource allocation planning.

While sales forecasting helps tremendously in planning for the future, additionally it helps us to analyze the performance of the organization by comparing the actual performance with the predicted or forecasted sales which act as a basis for evaluating the effectiveness of the sales strategies implemented for the financial year. In conclusion, sales forecasting helps in analyzing market demand which highly depends on customer or consumer behavior.

It is understood that there are a lot of factors influencing sales. One of the largest factors is called "inflation". In economics, Inflation is the gradual rise in the average price of goods and services over time in a given economy due to increased inflow of money for a very few goods available in the market. Increased government spending, a growing population, and increased demand from other nations can all contribute to this. Inflation is usually measured in terms of "inflation rate" which is expressed as a percentage change in the Consumer Price Index (CPI), which tracks changes in the prices of a variety of household-consumed products and services.

While there are some positive effects of moderate inflation like encouraging consumers to spend and invest on goods and services before prices increase further, high inflation can devalue a currency, discourage investment, and make consumers and businesses insecure. Therefore, it becomes very crucial to add the inflation factor in the process of sales forecasting. This paper hence focuses on including the parameter of inflation rate in the sales prediction model using the Dynamic Generalized Linear Model.

Inorder to regulate the inflation rates, every country has its own central bank. India's central bank is the "Reserve Bank of India (RBI)". By the preamble of RBI, RBI has the basic functions so as "to regulate the issue of Bank notes and keeping of reserves with a view to securing monetary stability in India and generally to operate the currency and credit system of the country to its advantage; to have a modern monetary policy framework to meet the challenge of an increasingly complex economy, to maintain price stability while keeping in mind the objective of growth".

The Reserve Bank of India manages inflation with monetary policies, such as increasing bank rates, repo rates, cash reserve ratios, dollar purchases, and managing money supply and credit availability. These actions reduce the market's money supply, decreasing demand and, consequently, prices. Nonetheless, the inefficiency of these technologies impedes growth and investment patterns.

Currently, the RBI has reiterated that the inflation rate target is 4% since the target provides flexibility. Besides, the annual consumer price inflation in India during February was 6.44%, and during January was 6.52%. Taking the current situation into consideration, the inflation rate has been set to 5% (taking an average of the target rate and the current rate) from the perspective of India.

## II. Issues in the Existing System

Existing sales prediction and forecasting models contain several flaws that diminish their accuracy and utility. Among these concerns are listed below:

Insufficient data integrity can result in inaccurate predictions and forecasts. This can include data that is missing, inconsistent, or obsolete. Some models may rely excessively on historical data, which can be problematic if market conditions or other factors have changed substantially since the data was collected.

Many models fail to adequately account for external factors that can affect sales, such as changes in the economy, the entry of new competitors, or changes in consumer behavior. Absence of transparency: Some models may be complicated and difficult to comprehend, making it difficult for businesses to evaluate their accuracy and dependability. Some models may not be adaptable enough to adjust to shifting market conditions or other variables that can affect sales.

Some models may only consider a limited number of variables, such as historical sales data or product inventory levels, and may overlook other crucial variables that can affect sales. Sales prediction and forecasting models may be susceptible to human error, such as inaccurate data entry or erroneous assumptions.

## III. Methodology

This paper utilises the Walmart Sales dataset to forecast sales. It is a well-known dataset containing weekly sales information for 45 Walmart stores located in various U.S. regions. The dataset includes information on store location, department, and classification, as well as various sales and promotional events that occurred between 2010 and 2012.

In predictive modelling and time series analysis, the Walmart sales dataset is often used as a benchmark. Researchers and practitioners utilise this dataset to develop and test new forecasting models and algorithms, evaluate the efficacy of various promotional strategies, and study the influence of external factors (such as holidays, weather, and economic indicators) on store sales.

The dataset is accessible to the public and can be downloaded from a variety of locations, including the Kaggle platform and the UCI Machine Learning Repository. The dataset consists of more than 140 million sales records and is very large, necessitating significant computational resources for processing and analysis.

The model used here for sales forecasting is called the "Dynamic Generalized Linear Model (DGLM)". It is a class of Bayesian state space models that can be used to analyze time series data.The central concept underlying DGLM is to characterise the time series as an evolving latent state process, and then to link this process to the observed data via a set of observation equations. Typically, the state process is represented by a linear Gaussian state space model in which the latent state at time 't' is a linear function of the previous state and a random Gaussian noise term. The observation equation is then utilised to map the latent state to the observed data at each time point, the mapping being determined by a link function that maps the linear combination of the latent state and observation noise to the mean of the observed data.

DGLM's flexibility in modelling a wide variety of time series data, including count data, binary data, and continuous data, is one of its advantages. This is accomplished by selecting a suitable link function for the observation equation. If the data are counts, for instance, a Poisson link function can be used, whereas if the data are binary, a logistic link function can be used.

Bayesian methods can be used to fit DGLM models, allowing for uncertainty quantification and model selection. Typically, the inference is conducted using Markov chain Monte Carlo (MCMC) methods, which permit sampling of the model parameters and latent states after the fact.

The state equation of the DGLM model is typically specified as a linear Gaussian state space model of the form:

$$y_t = F_t \Theta_t + v_t$$

$$\Theta_t = G_t \theta_{t-1} + w_t$$

where Y_t is the observed data at time t, _t is the latent state of the system at time t, _t and w_t are random Gaussian noise terms with mean zero and known variances, and F_t and G_t are matrices of coefficients that describe the relationship between the latent state and the observed data, and the evolution of the latent state over time, respectively.

## IV. Proposed Work

Adding an inflation factor to sales forecasting can help companies to better predict the impact of inflation on their sales revenue. Inflation refers to the general increase in prices of goods and services over time, which can impact consumer purchasing power and behavior.

By adding an inflation factor to their sales forecasting models, companies can more accurately predict the impact of inflation on their sales revenue. This can help them to better plan for the future, by identifying potential challenges or opportunities associated with inflation and adjusting their business strategies accordingly.

For example, if a company expects inflation to increase significantly in the coming year, they may adjust their pricing strategy or product offerings to better meet the needs of consumers who may be more price-sensitive due to inflation.

On the other hand, if a company expects inflation to remain low, they may be able to maintain their current pricing strategy or even raise prices without negatively impacting sales.

Overall, adding an inflation factor to sales forecasting can help companies to make more informed decisions about pricing, marketing, and other business strategies in order to optimize their sales revenue and profitability over time.

For companies of all sizes, forecasting and predicting sales may be a difficult undertaking. However, businesses may enhance their overall sales performance and produce more accurate sales estimates with the correct technology in place. Here is a suggested method for forecasting and predicting sales with inflation.

1) Data collection and analysis: Compiling information on previous sales performance is the first stage in creating a sales forecasting and prediction system. Information on sales volume, income, clientele demographics, and other details may be included. To find patterns and trends in this data, numerous statistical and machine learning approaches may be used for analysis.

2) Create predictive models: Predictive models may be created to estimate future sales performance based on the examination of historical sales data. These models may be founded on a variety of variables, such as consumer behaviour, market trends, and seasonal variations.

3) Include external data: Companies might include external data in their models to increase the precision of sales forecasts. This might contain data on rival activities, social media trends, and economic statistics.

4) Observe and correct: As with any predictive system, it's critical to keep an eye on the accuracy of sales projections and correct them as required. The models may need to be adjusted, new data sources may need to be added, or sales tactics may need to be changed.

5)Results dissemination: Disseminating the findings to important constituencies is the last phase in the sales prediction and forecasting method. Sales groups, managers, and investors may fall under this category. The alignment of expectations and enhancement of overall business performance may be achieved via clear and succinct communication of sales estimates and forecasts.

The below steps define how the DGLM Algorithm works: In the first stage, the DGLM model is specified by defining the linear predictor and the nonlinear link function. Typically, the linear predictor is a function of the state variables, which are considered to be governed by a particular dynamic system. The nonlinear link function is utilized to model the time series variance.

The next stage involves estimating the DGLM model's parameters. This can be accomplished through the use of maximum likelihood estimation or Bayesian methods. The coefficients of the linear predictor, the variance of the error term, and the parameters of the nonlinear link function must be estimated as parameters.

Following the estimation of the DGLM model's parameters, the state variables are filtered using a state-space model. This step is called "State filtering". The state variables characterize the fundamental process that produces the observed time series data. Utilizing the observed data to update the state variables and estimate their values at each time point constitutes filtering.

After the state variables have been filtered, the following step is to use them to forecast the future values of the time series. This is possible using the state-space model and the estimated DGLM model parameters.

In the concluding step, the performance of the DGLM model is evaluated by comparing its forecasts with the actual values of the time series. Various metrics, such as mean squared error, mean absolute error, and R-squared, can be used for this purpose. Based on the evaluation results, the model may be refined or altered.

Execution of the DGLM model involves specifying the model, estimating its parameters, filtering the state variables, forecasting future values, and evaluating the model's performance. This method permits a rigorous and systematic analysis of time series data using the DGLM model.

## V. Results

| Model | Accuracy |
|---|---|
| Sales Forecasting using Neural Networks | 100% |
| GBT Model | 98% |
| Decision Tree | 71% |
| GLM Model | 64% |

TABLE I
ACCURACY MEASUREMENTS OF EXISTING MODELS WITHOUT CONSIDERING THE FACTOR OF INFLATION

Neural networks for sales forecasting, we typically start by collecting historical sales data, such as daily, weekly or monthly sales figures over a period of time. We would then use this data to train a neural network model, which would learn to recognize patterns and relationships between different variables, such as sales volume, pricing, advertising, and other factors that might influence sales.

Once we have trained our neural network model, we can use it to make predictions about future sales. This might involve feeding in new data, such as current economic conditions or changes in marketing strategy, and using the model to generate forecasts about what sales might look like in the coming weeks, months, or years.

Sales forecasting using Gradient Boosted Trees (GBT) model involves using a machine learning technique to analyze historical sales data and make predictions about future sales. GBT is a popular ensemble learning technique that combines multiple decision trees to make accurate predictions.

To use GBT for sales forecasting, we typically start by collecting historical sales data, such as daily, weekly or monthly sales figures over a period of time. Then use this data to train a GBT model, which would learn to identify patterns and relationships between different variables, such as sales volume, pricing, advertising, and other factors that might influence sales.

Once we have trained our GBT model, we can use it to make predictions about future sales. This might involve feeding in new data, such as current economic conditions or changes in marketing strategy, and using the model to generate forecasts about what sales might look like in the coming weeks, months, or years.

To use decision trees for sales forecasting, we typically start by collecting historical sales data, such as daily, weekly, or monthly sales figures over a period of time. we would then use this data to train a decision tree model, which would learn to identify patterns and relationships between different variables, such as sales volume, pricing, advertising, and other factors that might influence sales.

Once we have trained your decision tree model, we can use it to make predictions about future sales. This might involve feeding in new data, such as current economic conditions or changes in marketing strategy, and using the model to generate forecasts about what sales might look like in the coming weeks, months, or years.

To use a GLM for sales forecasting, we typically start by collecting historical sales data, such as daily, weekly, or monthly sales figures over a period of time. We would then use this data to train a GLM model, which would learn to identify patterns and relationships between different variables, such as sales volume, pricing, advertising, and other factors that might influence sales.

The GLM model can be used to make predictions about future sales by inputting new data, such as current economic conditions or changes in marketing strategy. The model will generate forecasts about what sales might look like in the coming weeks, months, or years based on the relationships between the input variables and the response variable.

While the above results are done on models without considering the factor of inflation, this paper has its results which include the factor of inflation which is a start in the domain due to the current recession phase all over the world. The accuracy of the model tends to be around 91% which is better compared to 64% of the GLM Model.

## VI. Future Scope

Incorporating external factors that influence sales into sales forecasting algorithms can provide a more accurate and comprehensive picture of future sales trends. Incorporating external factors such as economic indicators, seasonality, competitor activity, and market trends can improve sales forecasting algorithms.

For instance, economic indicators such as GDP growth, unemployment rates, and consumer confidence can shed light on consumer spending patterns and market conditions as a whole. Holidays and seasonal trends can also influence sales patterns. Sales can also be affected by competitor activities, such as pricing strategies and marketing campaigns.

Businesses can alter their strategies and make more informed decisions regarding production, inventory, and pricing by incorporating external factors into their sales forecasting algorithms. This can help businesses remain competitive and capitalise on market trends, resulting in increased sales and revenue.

As businesses continue to acquire and analyse data from diverse sources, it is likely that the incorporation of external factors into sales forecasting algorithms will become more sophisticated and accurate. With the assistance of advanced analytics and artificial intelligence, businesses can gain a more in-depth understanding of the factors that influence their sales and modify their strategies accordingly.

Integrating an inflation factor into sales forecasting algorithms can improve their ability to determine future sales and revenue. By incorporating inflation into sales projections, businesses can anticipate the impact of rising prices on consumer behavior, modify their pricing strategies accordingly, and make more informed decisions regarding production and inventory.

Inflation can be incorporated into sales forecasting algorithms in a number of ways, including adjusting historical sales data for inflation, utilising economic indicators and forecasting models to predict future inflation rates, and integrating inflation assumptions into pricing and revenue projections.

As businesses become more sophisticated in their use of data and predictive analytics, it will likely become more common to incorporate inflation into sales forecasting algorithms. It can aid businesses in comprehending and adapting to shifting market conditions and consumer behavior, thereby enhancing their bottom line.

## VII. Conclusion

To make important decisions regarding future sales and income, businesses must use sales forecasting and prediction alongside inflation rates. Examining historical data, industry trends, and other pertinent factors enables businesses to

accurately predict future sales volumes, identify potential threats and opportunities, and develop successful sales strategies.

It is impossible to overstate the importance of accurate sales forecasting and prediction because it enables businesses to coordinate their operations, allocate their resources, and optimize their sales and marketing efforts. It aids businesses in identifying development opportunities, recognizing patterns, and drawing data-driven conclusions that may increase profits.

It is essential to remember that sales forecasting and prediction are not exact sciences and that there is always a margin of error. The accuracy of sales forecasts is impacted by the quality and quantity of data provided, the complexity of the industry, and any external factors that may affect sales volumes.

Despite their limitations, sales forecasting and prediction continue to be essential tools for businesses of all sizes. With the right data, analytical tools, and knowledge, businesses can make informed decisions regarding their sales strategy and optimize their operations to achieve greater long-term success.

## VIII. Discussion and Summary

There aren't papers that mainly focus on sales prediction considering inflation rates in the prediction model. The discussed papers here are normal forecasting models using different machine learning and deep learning models that exist in the industry.

This paper's approach of using the DGLM model for Walmart Sales Forecasting has been a useful instrument for sales forecasting because it can accommodate a wide variety of sales data, such as count data, binary data, and continuous data. More specific applications of the DGLM model for sales forecasting are listed below.

Seasonality is frequently present in sales data, which makes it difficult to predict future sales. By incorporating seasonal factors into the DGLM model, seasonality can be accounted for, resulting in more accurate sales forecasts.

The DGLM model is a dynamic model that can accommodate variations in the mean and standard deviation of the sales data over time. This makes it a more accurate model for sales forecasting, as sales patterns are frequently subject to change.

Incorporating external factors: The DGLM model can incorporate external factors that may influence sales, such as marketing campaigns, economic shifts, and consumer behavior changes. This enables more precise projections of future sales.

The DGLM model can generate accurate forecasts of future sales, which is useful for planning and making decisions. The model can be utilized to forecast sales at various time horizons, including monthly, quarterly, and annual.

The DGLM model permits the evaluation of the model's efficacy by comparing its sales forecasts to actual sales data. This permits the model's refinement and modification, which can lead to more precise forecasting in the future.

## Appendix A
## Data Visualization

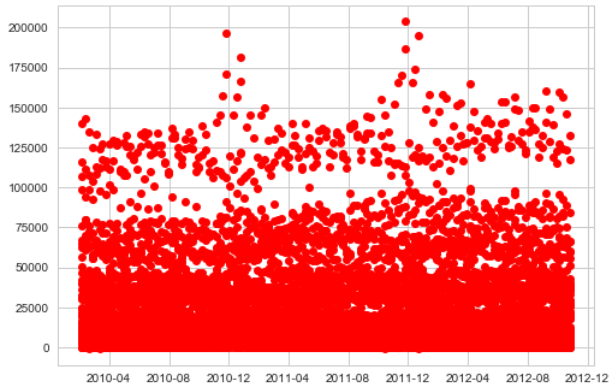


Fig. 1. Date Vs Overall Weekly Sales of all the stores, Shows the Weekly sales in every week represented by the starting date of the week.

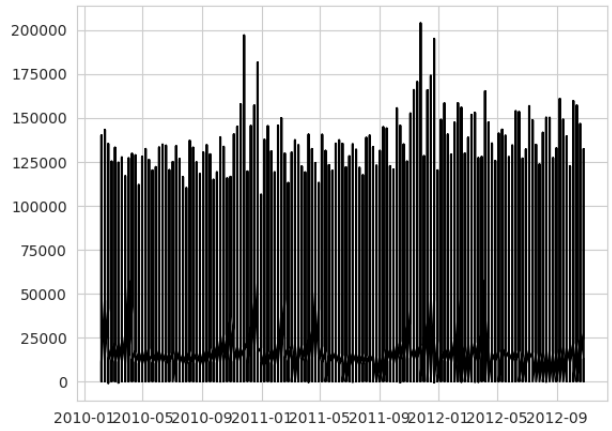


Fig. 2. Date Vs Overall Weekly Sales of all the stores Bar Graph, Shows the Weekly sales in every week represented by the starting date of the week

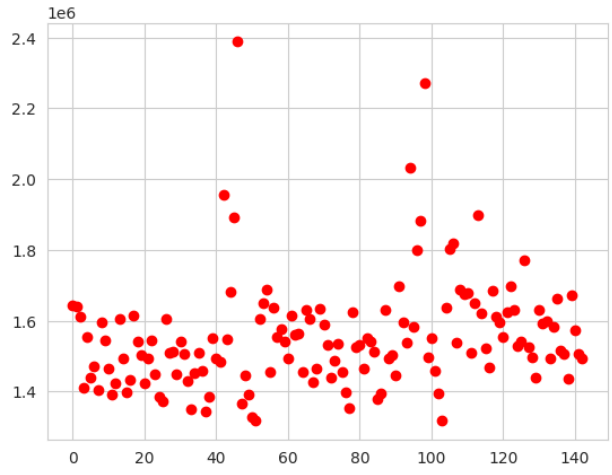


Fig. 3. Week Number Vs Weekly Sales of Store 1, Shows the sales generated in each week of the store 1 separately from the overall sales.

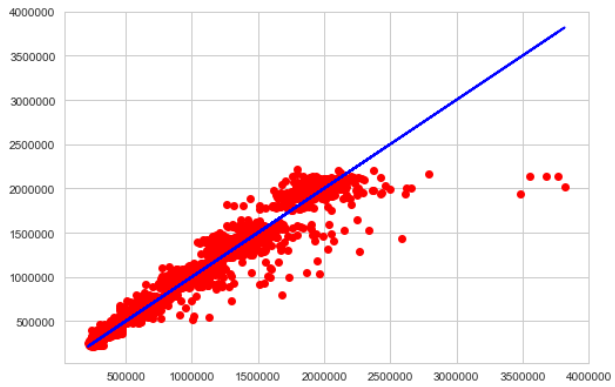


Fig. 4. Linear Prediction Graph of the Walmart Sales Dataset

## REFERENCES

[1] F. M. Thiesing and O. Vornberger, "Sales forecasting using neural networks," Proceedings of International Conference on Neural Networks (ICNN'97), Houston, TX, USA, 1997, pp. 2125-2128 vol.4, doi: 10.1109/ICNN.1997.614234

[2] Pavlyshenko BM. Machine-Learning Models for Sales Time Series Forecasting. Data. 2019; 4(1):15. https://doi.org/10.3390/data4010015

[3] Schmidt A, Kabir MWU, Hoque MT. Machine Learning Based Restaurant Sales Forecasting. Machine Learning and Knowledge Extraction. 2022; 4(1):105-130. https://doi.org/10.3390/make4010006

[4] X. dairu and Z. Shilong, "Machine Learning Model for Sales Forecasting by Using XGBoost," 2021 IEEE International Conference on Consumer Electronics and Computer Engineering (ICCECE), Guangzhou, China, 2021, pp. 480-483, doi: 10.1109/ICCECE51280.2021.9342304

[5] Cheriyan, S., Ibrahim, S., Mohanan, S., Treesa, S. (2018, August). Intelligent sales prediction using machine learning techniques. In 2018 International Conference on Computing, Electronics and Communications Engineering (iCCECE) (pp. 53-58). IEEE

[6] Wisesa, O., Adriansyah, A., Khalaf, O. I. (2020, September). Prediction analysis sales for corporate services telecommunications company using gradient boost algorithm. In 2020 2nd International Conference on Broadband Communications, Wireless Sensors and Powering (BCWSP) (pp. 101-106). IEEE

[7] M. Gao, W. Xu, H. Fu, M. Wang and X. Liang, "A Novel Forecasting Method for Large-Scale Sales Prediction Using Extreme Learning Machine," 2014 Seventh International Joint Conference on Computational Sciences and Optimization, Beijing, China, 2014, pp. 602-606, doi: 10.1109/CSO.2014.116

[8] Haselbeck, F., Killinger, J., Menrad, K., Hannus, T., Grimm, D. G. (2022). Machine learning outperforms classical forecasting on horticultural sales predictions. Machine Learning with Applications, 7, 100239

[9] Plott, C. R., Chen, K. Y. (2002). Information aggregation mechanisms: Concept, design and implementation for a sales forecasting problem

[10] Morwitz, V. G., Schmittlein, D. (1992). Using segmentation to improve sales forecasts based on purchase intent: Which "intenders" actually buy?. Journal of marketing research, 29(4), 391-405

[11] G. Srilakshmi amp; I. R. Praveen Joe (2023) "A-DQRBRL: attention based deep Q reinforcement battle royale learning model for sports video classification", The Imaging Science Journal, DOI: 10.1080/13682199.2023.2180022

[12] Praveen Joe I R, E M Malathy, S Aishwarya, R Akila, A Akshaya, "A Hybrid PSO-ACO Algorithm to Facilitate Software Project Scheduling", International Journal of e-Collaboration (IJeC), 2022, DOI: 10.4018/IJeC.304039

[13] Malathy, E. M., I. R. Praveen Joe, and P. Ajitha. quot;Miniaturized Dual-Band Metamaterial-Loaded Antenna for Heterogeneous Vehicular Communication Networks.quot; IETE Journal of Research (2021): 1-10.

[14] Joe, IR Praveen, and P. Varalakshmi. quot;A Two Phase Approach for Efficient Clustering of Web Services.quot; Computational Intelligence, Cyber Security and Computational Models. Springer, Singapore, 2016. 165-170.

[15] I. R. Praveen Joe amp; P. Varalakshmi (2019) A Multilayered Clustering Framework to build a Service Portfolio using Swarm-based algorithms, Automatika, 60:3, 294-304, DOI: 10.1080/00051144.2019.1590951

[16] Praveen Joe, I.R. and Varalakshmi, P. (2019) quot;An Analysis on Web-Service-Generated Data to Facilitate Service Retrieval,quot; Applied Mathematics amp; Information Sciences: Vol. 13: Issue 1, pp:47-55 DOI: http://dx.doi.org/10.18576/amis/130107

[17] IRP Joe, G Washington , "Art network—a solution for effective warranty management" - CURIE J, 2008

[18] Moorthy Agoramoorthy, Irudayaraj Praveen Joe, "Hybrid cuckoo–red deer algorithm for multiobjective localization strategy in wireless sensor network", International Journal of Communication Systems, https://doi.org/10.1002/dac.5042, December 2021.

[19] Joe, Praveen. quot;IR and Varalakshmi, P,A Survey on Neural Network Models for Data Analysis.quot; ARPN Journal of Engineering and Applied Sciences 10.11 (2015): 4872-4876.

[20] V. Dhanasekar, Y. Preethi, V. S, P. J. I. R and B. P. M, quot;A Chatbot to promote Students Mental Health through Emotion Recognition,quot; 2021 Third International Conference on Inventive Research in Computing Applications (ICIRCA), 2021, pp. 1412-1416, doi: 10.1109/ICIRCA51532.2021.9544838.