DYNAMIC SALES COMMISSION PREDICTION USING MACHINE LEARNING

V Karthick, Associate Professor
Department of CSE
Rajalakshmi Engineering College
Chennai, India
vkarthick86@gmail.com

Yaashish G , UG Student
Department of CSE
Rajalakshmi Engineering College
Chennai, India
210701317@rajalakshmi.edu.in

Shiyaam Prasad V, UG Student
Department of CSE
Rajalakshmi Engineering College
Chennai, India
210701321@rajalakshmi.edu.in

ABSTRACT -To achieve better revenue performance, to motivate the sales team and to increase the general performance of the business, it is crucial to match sales commission in the highly competitive market of today. However, the traditional commission structures often fail to adequately accommodate the dynamics of product demands that lead to inefficiencies, lost opportunities, and even injustice among the salespeople. To solve these issues, the proposed approach utilizes artificial intelligence and machine learning techniques. [13] is a paper which seeks to improve dynamic sales commission prediction using semi-supervised learning technique by combining both limited labeled data and large amounts of unlabeled data to achieve high accuracy and less annotation costs. It adjusts commission rates based on the changes in the product demand by using the historical sales data, real-time market information and analytical models. This paper [19] combines decision analysis and predictive modeling and uses linear regression for accurate sales prediction for dynamic sales commission rates. Thus, the strategy is effective in ensuring that sales executives are motivated to sell the products with slow moving inventory. thus enhancing the possibility of stock clearance while minimizing holding costs for inventories and at the same time maximizing revenue prospects.

Keywords -Linear regression, Dynamic,Mean absolute error,Root mean squared error.

I. INTRODUCTION

In retail sales, sales commissions are crucial for driving profits and motivating sales teams. The traditional way of calculating commissions has issues adapting to market changes and customer preferences. The proposed solution, the Dynamic Sales Commission Generator, uses advanced technologies like machine learning and web development to automate calculations and provide strategic insights. This dynamic sales commission predictor automates fee calculations, predicts income commissions, and gives actionable insights for strategic selection-making.

Previous studies in sales commission optimization and dynamic pricing have laid a strong basis for grasping the complexities of fee systems and their impact on income overall performance. Research has shown the importance of optimizing sales commissions and dynamic pricing to boost revenue. However, there is a lack of information on leveraging predictive analytics for dynamic commission optimization. This study aims to fill this gap by developing a framework for real-time commission prediction using machine learning. Some of the reference papers are listed below. [12] is a paper listing the application of semi-supervised learning to enhance dynamic sales commission prediction by leveraging both limited labeled data and abundant unlabeled data for improved accuracy. [13] is another paper regarding the machine learning techniques that explains linear regression with

least square and least absolute deviation methods to geographical data, demonstrating how imprecise coefficients can improve dynamic sales commission prediction by handling uncertainty in the data. These two papers were mainly considered as references for choosing the machine learning algorithm for this dynamic sales commission prediction project.

1] "Paper: How well does predictive analytics tools work in sales for customer churn prediction"; from which I extracted only a bit relevant to this use-case. They verified this model with more than 9.7 million transactions conducted in a real-world retail environment of business-to-business company, which enabled them to accurately predict individual purchasing behavior Whether this model works or not is a joint effect of customer characteristics like probability to churn, and tenure etc, and salespersons' slant towards technology (iot in this context) and skill set. And they discovered that "setting more realistic expectations about the tool's accuracy" was somewhat effective, but only under some circumstances.

By the help of this paper [2], it is argued that you can divide the salespeople's compensations according to their acceptance of risk. They know that low risk taking propensity salespeople will ensure that they do more than is necessary to keep existing customers while such a situation might not be the case with high risk taking propensity salespeople. Besides, they describe other payment techniques aimed at keeping motivation of both types and at not relying too much on the accomplishments achieved earlier. Therefore, it assists in helping the firms to design fair and effective compensation methods in relation to their sales personnel.

In this paper [3], the authors provide explanations that detail how accurately the sales impact overall costs. Analyzing data from EU surveys together with a machine learning tool, the xgboost, it is shown that worse sales changes that are not forecasted predict even greater variations in costs in contrast with better predictable sales changes. They also learn that it is very difficult for them to disguise costs especially at a time when sales are least as expected. This paper contributes to the knowledge and practice of managers as well as researchers in organisation by revealing the consequences of sales prediction on cost management. Also, by the inclusion of xgboost, it provides guidance for medium-sized firms that are limited in data access.

As mentioned in Paper [4], the study was to determine optimal commission rates of a salesman who is dealing in multiple products. This must reveal how these rates should be determined on the basis of the presumption that sales grow in a linear fashion with the time devoted to selling a product. However, it also tries to discover what happens when the sales are not growing in a linear-like fashion, explaining how this influences the ideal commission rate to represent mutually beneficial conditions for the company and the salesman.

To my unfairing, this review paper [5] presents a view on the application of artificial intelligence in the prediction of sales commission. Thus, the evaluation of the models is done by analysing them intensively in terms of different criteria. Extensive literature review which focused on 120 research papers published between 2008 and 2021 are reviewed. As highlighted by the authors in their respective research, crimes and spatial are the most used to address the issue of crimes. Each of the 120 researched papers was observed concerning the type of ML models that was used as shown in the section on the distribution of ML models, it was observed that supervised learning models formed the biggest group contributing 31% while models that incorporated both the supervised learning models and the unsupervised learning models contributed 22% while the sole unsupervised learning models contributed contributed 10%.

This study [6] is intended to make people conscious about the significance of sales prediction and sales forecasting for the betterment of sales division. Incorporate and advance on prior studies, which fail to accurately capture learning models, this study employs more precise models such as SVM, XGBoost, KNN, and ARIMA models to fit the crime data. The work

suggests that these predictive models can help support sales and marketing.

This work [7] focuses on how the assignment of territories and financial remunerations of the salespeople impacts the profit of the company. They also identify that paying salespeople based on the group they are with, even if they are in different territories that have different levels of sales increases the money made by the company. To sum it up, in larger sales teams, it is more advantageous to assign equitable quotas that are to be paid with the group. They also examine different types of payments and identify that sometimes it is more effective to pay on the specific sales performance while in other cases, it is possible to use contests.

This review paper [8] is mostly dedicated to the basics and the algorithms of machine learning in simple terms. It introduces core concepts and depicts how they can be incorporated to form real-world algorithms. It also covers more general topics such as computational complexity and several approaches to performing the algorithms. It is an amazing course that targets matrix, statistics, computer science, mathematics, engineering, and any student who wants to learn about machine learning.

This paper [9] proposes the use of big data machine learning algorithms for crime forecast based on the historical sales commission data of a super market in China of between 2015 and 2018. It concludes that LSTM model yields a much higher accuracy compared to other algorithms such as KNN, SVM, RFM and the results show that, the integration of environmental factors has shown better performance over the ordinary LSTM model

This particular study [10] aims to identify the most effective method of motivating the target sales teams under a dynamic pay model without the limitations as to contracts. I also explore when traditional pay structures, such as fixed or predictable wages, quotas or commission-based programs are most effective. For instance, fixed wages are ideal for low sales turnover while commissioning is more appropriate during high turnover. It also incorporates situations such as variation in demand, and consequently, preference variability.

This paper [11] proposes a new forecast method of product sales based on the Bass/Norton model incorporating sentiment analysis from an online review. The study looks to improve sales forecasts through the use of the Naive Bayes algorithm to analyse sentiments within the reviews. In doing so it applies this method to actual data from the automotive industry and demonstrates its superiority to previous techniques. The combination of both helps to predict the sales of products and gain a better insight in to consumer behavior.

By analyzing historical sales data and market trends, this study seeks to create models for accurately predicting sales commissions quickly. The end goal is to help retail organizations improve their commission structures and sales performance strategically.

II MATERIALS AND METHODS

For this project on Dynamic Sales Commission prediction, we have adopted a systematic approach involving data collection, data preprocessing, model selection, training, and evaluation. The data set used here contains columns for sales amount, product demand level, and other relevant details. The data is preprocessed using Pandas to handle missing values and then extracts the required input features for the prediction task. The preprocessed data is then used to train a Linear Regression model using the Scikit-learn library. So this model learns the relationship between the input feature and the target (sales commission). The model then predicts the sales commissions for the given sales amount and product demand level from the uploaded dataset. This predicted sales commission is then displayed to the user in the front end as JSON responses. The accuracy and performance of the model is evaluated using cross-validation techniques. This evaluation is done during the model training phase. Finally, the model's predicted sales commission is

validated against the actual sales commission to assess the model's prediction effectiveness.

Hardware requirements of the project include

- Operating System (Windows/Linux)

- RAM (4 GB Minimum)

- Processor: Multi-core CPU

- Graphics Processing Unit (GPU)

- Secondary Storage (256 GB Minimum)

Software requirements include

- Internet browser (Chrome/Edge/Mozilla Firefox)

- Internet connection

- Programming Language and frameworks

    (Python, Flask, JavaScript, HTML/CSS, Bootstrap)

- Pandas, Numpy

- Scikit-learn, Joblib

III EXISTING SYSTEM

The existing sales commission prediction system has several hurdles that affect its effectiveness. A major limitation is its lack of adaptability. obstacles are created by the rigidity if the system for the organization in optimizing sales strategies and to align commission rates with business objectives. moreover, the manual sales commission calculation system is error-prone and may result in making mistakes in commission calculation and create possible disagreement among the sales employees. These mistakes not only breaks rust within the sales community but also adds more burdens to the administration.[2] paper highlights the importance if sales prediction accuracy in measuring the cost compacts.the previous studies has examined the effects of expectation on cost behaviour, focusing on predicted and un predicted sales changes. mainly the findings of the paper indicates that commission increases higher for unpredicted sales compared to predicted sales.In Addition, the manual process of commission calculation wastes valuable time and resources, reduces its effectiveness, and redirects the focus away from the main strategic sales activities. The lack of real-time data analysis and insights limits the capabilities of the organization to make informed decision about commission rates and sales techniques. Relying only on past data can limit the ability of the organization to adapt to market change and growing sales fashions. Additionally [14] paper findings shows that the existing system rely on ad hoc restrictions which limit there flexibility and there is a lack of integration between quota based bonuses and commission schemes to handle complex market conditions. Even the paper [15] The system used does not support or incorporate dynamic changes in compensation to reflect changing sales dynamic over time.Finally, the fixed sales commission rated drastically reduces individual sales contribution causing differences in motivation among the sales organization and demoralizing top performers.Which would lead to the failure of the organization. Removing these issues requires a dynamic sales commission approach using machine learning algorithms and real-time analysis.Such a system can automate sales calculation and predict the sales commission thereby enhancing the sales performance effectively.

IV PROPOSED SYSTEM

The proposed system offers a significant enhancement in sales commission management by addressing numerous challenges faced by the current exsisting system. firstly, it generates dynamic commission rates which are calculated in real-time based on various

factors such as current sales data, market demand, and predictive analysis. The dynamic approach used by the dynamic sales commission predictor ensures to adjust the commission rates automatically to optimize sales performance and thereby increasing the revenue and providing a competitive advantage in the fast paced markets. Secondly, the system automates commission calculation using advanced machine learning algorithms. This system not only enhances the accuracy of the commission prediction but also save time for the sales employees by reducing manual efforts and errors associated with the exsisting model. Thereby allowing the sales employees to concentrate and focus more on the strategic task for sales. moreover the real time analysis of the model enhances the sales organization by providing insight of the sales performance, customer behavior and market trends. These insights help in empowering the decision makers to make effective and efficient decisions and adjust commission rates that align with the business goals. the user-friendly web interface enables smooth and interactive web pages allowing users to easily upload the required dataset and to visualize sales performance through understandable and optimal charts and to obtain accurate predictions for the sales commission. By utilizing the modern technologies such as Flask, HTML, CSS, Bootstrap, and Javascript like Chart.js, the system provides a smooth and interactive user experience ,ensuring high success rates among users. Lastly by paving way for the commission management process, ensuring transparency in commission allocation, and providing data driven insights thereby enhancing sales performance leading to organizational success.
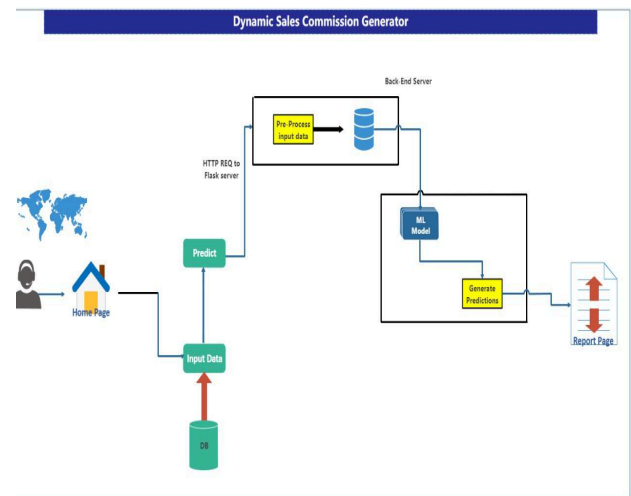

Figure 1. Architecture diagram

V METHODOLOGY

The algorithm used in the provided coding for the Dynamic Sales Commission Generator system is Linear Regression. Linear Regression is a supervised machine learning algorithm used for predicting a continuous output variable based on one or more input variables. In the context of the system, Linear Regression is used to predict sales commissions based on two input features: sales amount and product demand level.

A. Data preparation :

The sales data is a csv file which is uploaded by the user. The data set used here contains columns for sales amount, product demand level, and other relevant details. The data is preprocessed using Pandas to handle missing values and then extracts the required input features for the prediction task.

B. Model Training:

The preprocessed data is then used to train a Linear Regression model using the Scikit-learn library. So this model learns the relationship between the input feature and the target (sales commission).

All these samples can be continuously captured or any faulty image taken in between can be

clear according to a last in first out fashion.

C. Model Prediction:

Once the dataset is uploaded, The model then predicts the sales commissions for the given sales amount and product demand level from the uploaded dataset. This predicted sales commission is then displayed to the user in the front end as JSON responses.

D. Evaluation and Validation:

The accuracy and performance of the model is evaluated using cross-validation techniques. This evaluation is done during the model training phase. Finally, the model's predicted sales commission is

validated against the actual sales commission to assess the model's prediction effectiveness.
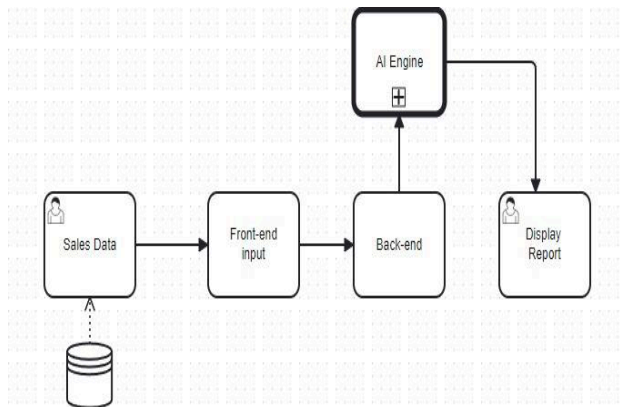

Figure 2. System flow diagram

VI RESULT

Thus the dynamic sales commission prediction model provides promising results. It successfully predicts accurate and efficient sales commission dynamically based in real time sales data , market demand, and predictive analysis. This model shows high accuracy as demonstrated by evaluation methods such as mean absolute errors (MAE), root mean squared error

(RMSE) and coefficient of determination (R-squared).Overall, the implementation of the dynamic sales commission prediction system enhances efficiency, transparency, and fairness in commission structures, resulting in benefits such as increased sales performance, maximization of revenues, and improved stakeholder satisfaction.

The metrics that we focussed on for prediction of sales commission achieved a greater accuracy with a quicker and yet stable response time.

VII DISCUSSION

The Cross-validation scores did offer ranging from approximately 0.739 to 0.946, with an average cross-validation rating of about zero.858. Generally, higher pass-validation scores indicating better performance, so those rankings do show that your version sort of performs gradually throughout numerous subsets of the training records. Concerning the evaluation metrics at the test set:

The intended squared mistakes (MSE) of about 424.27 does imply an average squared distinction between the predicted and real income commissions. Lower values of MSE do indicate sort of ok version performance.

Sort of R-squared score of approximately zero.931 does constitute form of the percentage of variance inside the target variable (sales commissions) this is sort of defined by the independent variables (income quantity and product call for degree.

In precise, the go-validation ratings and assessment metrics (MSE and R-squared) type of display that your model kind of does quite truly nicely in predicting income commissions based on the given features. However, it is kind of essential to take into account the precise context and requirements of your software to determine if these performance metrics without a doubt meet your expectations..
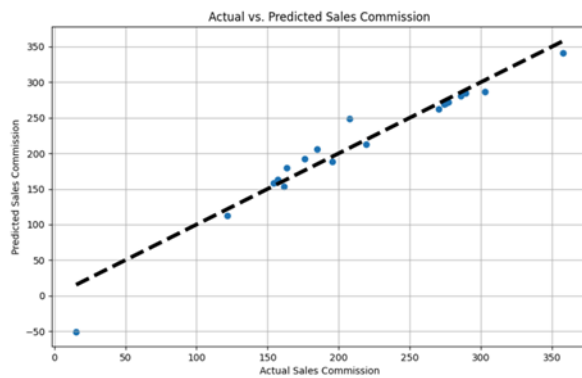
Figure 3. Comparison of actual sales commission and predicted sales commission

VIII CONCLUSION

Thus, the dynamic Sales Commission Generator project uses machine learning algorithms, particularly linear regression within a user friendly web application built using Javascript technologies. This solution effectively predicts sales commission based on the sales amounts and product demand levels, thereby paving way for the complex and time consuming process of commission calculation. Additionally provides a platform for users to upload dataset, receiving prediction and visualizing insights through interactive and understandable charts.It offers the sales business a significant tool to optimize sales strategies, motivate salesperson through incentives and maximize profitability. overall, This system has a potential to transform commission calculation process across various industries inorder to achieve today's competitive market.

REFERENCES

[1] Johannes Habel, Sascha Alavi and Nicholas Heinitz -"Effective Implementation of Predictive Sales Analytics," published on - December 30, 2022

[2] Olivier Rubel, Ashutosh Prasad "Dynamic Incentives in Sales Force Compensation," - July 10, 2016.

[3] Kira Hoffmann, Matthias D. Mahlendorf, Kim Pettersson "The Cost Impact of Sales Prediction Accuracy",- February 9, 2021.

[4] Paul Berger -"On Setting Optimal Sales Commissions "-01 June 2007.

[5] Fabio Caldieraro, Anne T. Coughlan, "Optimal Sales Force Diversification and Group Incentive Payments" – 19 june, 2009.

[6] Songjun Luo, Ph.D - "Does Your Sales Incentive Plan Pay for Performance?" - January 2003.

[7]S.Shalev-shwartz And S. Ben-david, Understanding Machine Learning: From Theory To Algorithms: Cambridge University Press, 2014.

[8]K. P. Murphy, Machine Learning: A Probabilistic Perspective: Mit Press, 2012.

[9] P. Domingos, "A Few Useful Things To Know About Machine Learning," Communications Of The Acm, Vol. 55, Pp. 78-87, 2012.

[10]Bargarai, F., Abdulazeez, A., Tiryaki, V., & Zeebaree,D." Management Of Wireless Communication Systems Using Artificial Intelligence-based Software Defined Radio."-2020

[11] B. Akgün And Ş. G. Öğüdücü, "Streaming Linear Regression On Spark Mllib And Moa," In Proceedings Of The 2015 Ieee/Acm International Conference On Advances In Social Networks Analysis And Mining 2015, 2015, Pp. 1244-1247.

[12]M. H. Dehghan, F. Hamidi, And M. Salajegheh, "Study Of Linear Regression Based On Least Squares And Fuzzy Least Absolute Deviations And Its Application In Geography," In 2015 4th Iranian Joint Congress On Fuzzy And Intelligent Systems (Cfis), 2015, Pp. 1-6.

[13] D. M. Abdulqader, A. M. Abdulazeez, And D. Q.

Zeebaree, "Machine Learning Supervised Algorithms Of Gene Selection: A Review," Machine Learning, Vol. 62, 2020.

[14] Matthias Kräkel , Anja Schöttner -"Optimal sales force compensation" - June 2016.

[15]Zhi-Ping Fan, Yu-Jie Che, Zhen-Yu Chen -"Product sales forecasting using online reviews and historical sales data: A method combining sentiment analysis"-May 2017.

[16]Matthew J. Schneider, Sachin Gupta -"Forecasting sales of new and existing products using consumer reviews: A random projections approach"-June 2016.

[17]Caroline E. W. Glackin, Murat Adivar-"Using the power of machine learning in sales research: process and potential"-July 9,2021.

[18]Akshay Krishna; Akhilesh V; Animikh Aich; Chetana Hegde-"Sales-forecasting of Retail Stores using Machine Learning Techniques"- December 22,2018.

[19]Sunitha Cheriyan, Shaniba Ibrahim, Saju Mohanan, Susan Treesa-"Intelligent sales prediction using machine learning techniques"-Agust 17,2018

[20]Sanjog Misra & Harikesh S. Nair-"A structural model of sales-force compensation dynamics: Estimation and field implementation"- 12 March, 2011.