DYNAMIC SALES COMMISSION PREDICTION

MINI PROJECT REPORT

Submitted by

Yaashish G	**210701317**
Shiyaam Prasad V	**210701321**

in partial fulfillment for the award of the degree of

BACHELOR OF ENGINEERING

in

COMPUTER SCIENCE AND ENGINEERING

RAJALAKSHMI ENGINEERING COLLEGE, CHENNAI

ANNA UNIVERSITY:: CHENNAI 600 025

APRIL 2024

RAJALAKSHMI ENGINEERING COLLEGE,

CHENNAI

BONAFIDE CERTIFICATE

Certified that this Report titled "**Identification of Patterns in Crime Records using Ensemble learning Approach**" is the bonafide work of **"YAASHISH G(210701317) and SHIYAAM PRASAD V(210701321) "** who carried out the work under my supervision. Certified further that to the best of my knowledge the work reported herein does not form part of any other thesis or dissertation on the basis of which a degree or award was conferred on an earlier occasion on this or any other candidate.

SIGNATURE

Karthick V
Assistant Professor,
Department of Computer Science and Engineering,
Rajalakshmi Engineering College,
Chennai – 602015

Submitted to Mini Project Viva-Voce Examination held on _____

Internal Examiner **External Examiner**

ABSTRACT

Initially For better revenue growth, keeping the sales team motivated and enhancing overall business performance, it is imperative to efficiently distribute sales incentives in today's highly competitive market. However, traditional commission structures are often not able to handle the changing nature of product demand adequately that result inefficiency,missed opportunities and potential inequalities in remuneration among sales teams.To address these problems, the suggested approach uses artificial intelligence and machine learning technologies. [13] is a paper which investigates 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 and reduced annotation effort. It adjusts commission rates dynamically depending on the fluctuations in product demand through leveraging historical sales data, real-time market insights, and advanced algorithms. This paper [19] integrates decision analysis and predictive modeling, utilizing Linear regression for highly accurate sales forecasts, to optimize dynamic sales commission predictions.In this manner, the strategy keeps sales executives motivated to sell products with slow moving inventory. thus optimizing stock clearance while reducing holding cost for inventories and maximizing revenue potentials.

ACKNOWLEDGEMENT

Initially we thank the Almighty for being with us through every walk of our life and showering his blessings through the endeavour to put forth this report. Our sincere thanks to our Chairman **Mr. S.MEGANATHAN, B.E, F.I.E.**, our Vice Chairman **Mr. ABHAY SHANKAR MEGANATHAN, B.E., M.S.,** and our respected Chairperson **Dr. (Mrs.) THANGAM MEGANATHAN**, **Ph.D.,** for providing us with the requisite infrastructure and sincere endeavoring in educating us in their premier institution.

Our sincere thanks to **Dr. S.N. MURUGESAN, M.E., Ph.D.,** our beloved Principal for his kind support and facilities provided to complete our work in time.

We express our sincere thanks to **Dr. P. KUMAR, Ph.D.**, Professor and Head of the Department of Computer Science and Engineering for his guidance and encouragement throughout the project work. We convey our sincere and deepest gratitude to our internal guide, **Karthick V** Professor, Department of Computer Science and Engineering. Rajalakshmi Engineering College for his valuable guidance throughout the course of the project.

SHIYAAM PRASAD V-210701321
YAASHISH G-210701317

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS…

FLASK	A micro web framework written in Python
HTML	Hypertext Markup Language
MSE	Mean Squared Error
NumPy	Numerical Python
JSON	Java Script Object Notation
CSV	Comma-Seperated Values
B2C	Business To Consumer

CHAPTER 1
INTRODUCTION

1.1 GENERAL

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.

1.2 OBJECTIVE

The main aim of this project is to analyse 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.

1.3 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.

1.4 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.

CHAPTER 2

LITERATURE SURVEY …

This paper [1] discusses how well predictive analytics tools work in sales, especially for predicting customer churn. They used data from 9.7 million transactions at a business-to-business company to create a model and tested it in a real-world setting. The success of this model depended on customer factors like their likelihood of leaving and past spending, as well as salesperson factors like their views on technology and sales skills. They also found that setting realistic expectations about the tool's accuracy helped, but only in certain situations. This study gives useful advice for managers on how to get the most out of their investment in analytics.

This paper [2] suggest ways to pay salespeople based on their risk tolerance. They find that less risk-averse salespeople work harder to keep existing customers, while more risk-averse ones may not. They recommend different payment plans to keep both types motivated and ensure they don't rely too much on past success. This helps companies design fair and effective compensation strategies for their sales teams.

This paper [3]describes describe how accurately predicting sales affects costs. Using data from EU surveys and a machine learning tool called xgboost, they find that unexpected sales increases lead to bigger cost jumps compared to predicted increases. They also discover that costs are harder to adjust when sales unexpectedly drop. This study helps managers and researchers understand the impact of sales prediction on cost management. Plus, by using xgboost, it offers insights for medium-sized firms with limited data access.

Paper [4] describes the study to decide on commission rates for a salesman selling different products. It talks about the best way to set these rates, assuming that sales increase predictably with the time spent selling a product. But it also explores what happens when sales don't always follow a predictable pattern, showing how this affects the ideal commission rates for both the company and the salesman.

This review paper [5] gives a take on Artificial Intelligence being used in the field of sales commission prediction. By intensively analysing various criteria, the models are evaluated. The research has concluded that crimes and spatial are the most applied categories in analysing crimes. The various ML models used across the 120 research papers were noted and supervised learning models were found to be the major contributors with 31% while a combination of supervised and unsupervised learning models contributed with 22% and unsupervised learning models alone contributed with 10%.

This study [6] draws attention towards the importance of sales prediction and forecasting to enhance sales division progress. As an improvement to the existing studies which lack accuracy on learning models, this study uses various machine learning algorithms like SVM, XGBoost, KNN and ARIMA model to better fit the crime data. The study concludes that these predictive models can aid support sales and marketing .

This study [7] concentrates on the way salespeople are assigned territories and paid affects a company's profits. They find that when salespeople work in territories with different sales outcomes and get paid as a group, it helps the company make more money, even if some territories sell less. In bigger sales teams, it's better to have a fair distribution of salespeople with group payments. They also compare different payment methods and find that sometimes it's better to pay based on sales performance, while other times it's better to use contests.

This review paper [8] focuses on the machine learning basics and algorithms in an easy-to-understand way. It covers fundamental ideas and explains how they're used to create practical algorithms. It also talks about advanced topics like computational complexity and different algorithmic approaches. Designed for students with backgrounds in various fields like statistics, computer science, math, and engineering, it's accessible to anyone interested in learning about machine learning.

This paper [9] compares machine learning algorithms for crime prediction using historical data of sales commission from a supermarkets in China between 2015 and 2018. It finds that LSTM model outperforms other algorithms like KNN, SVM, Random Forest and that incorporating environmental factors improves prediction accuracy compared to the model Thus the paper concludes that sales prediction techniques should use both environmental factors and historical sales to maximize the accuracy.

This study [10] explores the best ways to pay sales teams using a dynamic model without restricting contract options. It investigates when common compensation plans like fixed salaries, quota-based bonuses, or commissions work best. For example, fixed salaries are good for low sales revenues, while commissions are better for high revenues. It also looks at uncertain situations like demand and preference variability. The findings help understand real-world sales compensation and offer practical insights.

This paper [11] suggests a new way to predict product sales by combining the Bass/Norton model with sentiment analysis of online reviews. By using the Naive Bayes algorithm to understand sentiments in reviews, the study aims to make sales forecasts more accurate. It tests this method using real automotive industry data and finds that it works better than traditional approaches. This combination helps understand consumer behavior better and improves predictions of product sales.

This paper [12] aims to predict product sales using both numerical data and customer reviews. Most methods for this are complicated and costly. Instead, we propose a simpler approach called bag-of-words, but it can lead to too many predictors, causing problems. To fix this, we suggest using random projections to make the method more efficient. We tested our approach on Amazon.com sales data for tablets, and it performed better than other models. Plus, it's easy to use for different products and can handle large datasets.

Linear regression is one of the best practices that can be used to find out patterns and relationships within the dataset. This paper deals with [13] analysing out the pros and cons of each technique. This technique is mainly used in prediction techniques

This paper [14] explores how using data mining and machine learning can enhance sales research. It makes sales research easier for researchers and managers by improving accuracy and efficiency. Machine learning can predict salesforce performance more accurately by analyzing behavior, attitudes, and demographics. The study also suggests future research ideas and applications of machine learning in various sales contexts, such as B2C and B2B markets for companies and independent sales teams."

This paper explores how intelligent decision systems can improve sales forecasting using data mining techniques. Paper[15] analyzes sales data and different forecasting methods, we find that the Gradient Boost Algorithm stands out for its accuracy and reliability in predicting future sales trends.

CHAPTER 3
SYSTEM DESIGN

3.1 DEVELOPMENT ENVIRONMENT

3.1.1 HARDWARE SPECIFICATIONS

This project uses minimal hardware but in order to run the project efficiently without any lack of user experience, the following specifications are recommended

Table 3.1.1 Hardware Specifications

PROCESSOR	Intel Core i5
RAM	4GB or above (DDR4 RAM)
GPU	Intel Integrated Graphics
HARD DISK	6GB
PROCESSOR FREQUENCY	1.5 GHz or above

3.1.2 SOFTWARE SPECIFICATIONS

The software specifications in order to execute the project has been listed down in the below table. The requirements in terms of the software that needs to be pre-installed and the languages needed to develop the project has been listed out below.

Table 3.1.2 Software Specifications

FRONT END	HTML, CSS, Bootstrap, JavaScript
BACK END	Python, Flask, Joblib
FRAMEWORKS	Flask, Bootstrap
SOFTWARES TOOLS USED	Pandas,NumPy,Scikit-learn

3.2 SYSTEM DESIGN

3.2.1 ARCHITECTURE DIAGRAM

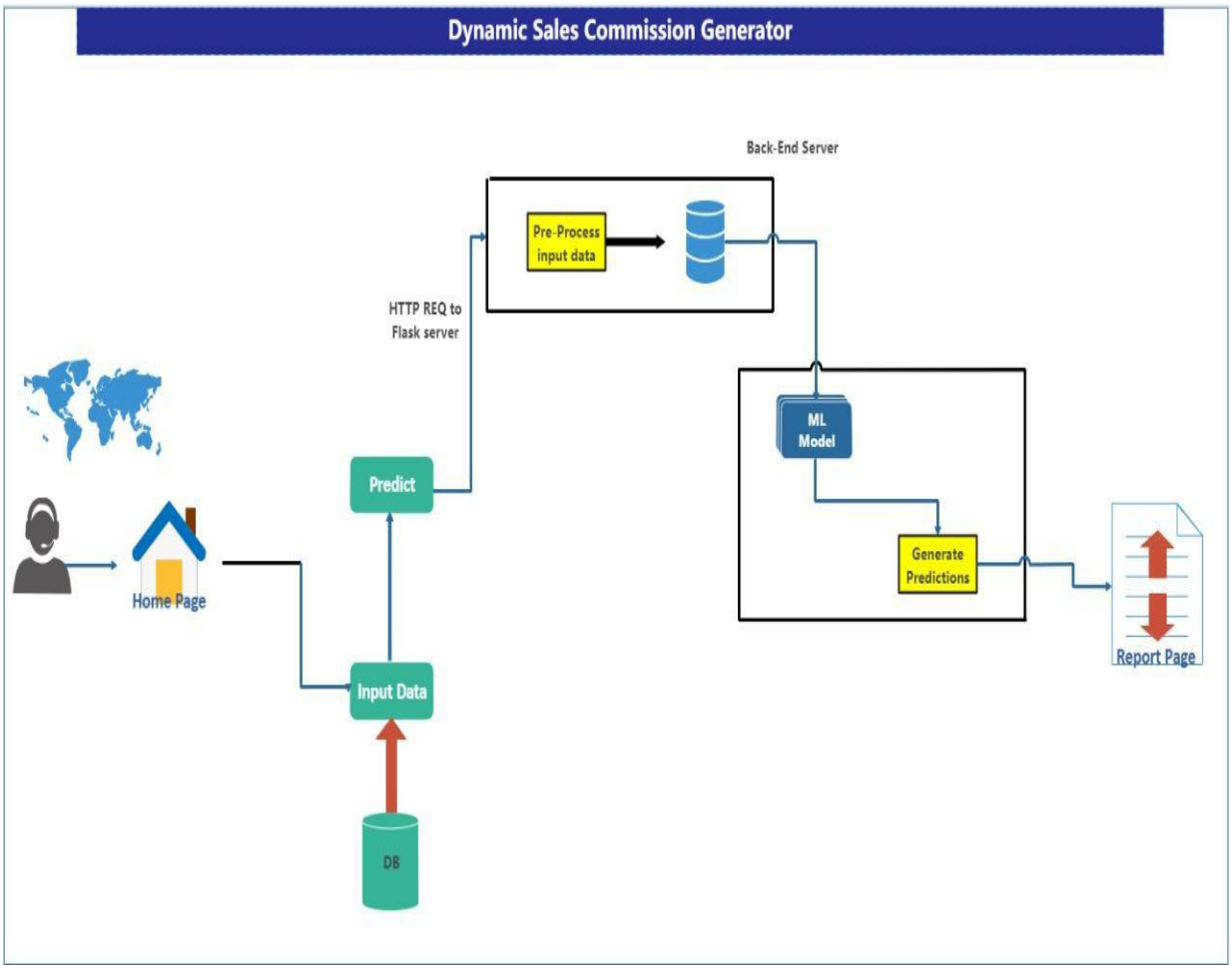

Fig 3.2.1 Architecture Diagram

CHAPTER 4
PROJECT DESCRIPTION

4.1 MODULE DESCRIPTION

4.1.1 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.

4.1.2MODEL 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.

4.1.3MODEL 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.

4.1.4EVALUATION 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.

CHAPTER 5
IMPLEMENTATION AND RESULTS

5.1 IMPLEMENTATION

The implementation of the Dynamic Sales Commission Prediction project involved several key steps. Firstly, we collected historical sales data, including sales amounts, product demand levels, and other relevant details. This data was then preprocessed using Pandas to handle missing values and extract the required input features for the prediction task. Subsequently, we selected the Linear Regression algorithm from the Scikit-learn library to train our predictive model. This model learned the relationship between the input features and the target variable, which in this case was the sales commission. During the model training phase, we employed cross-validation techniques to evaluate its accuracy and performance. Once trained, the model was deployed to predict sales commissions based on given sales amounts and product demand levels. These predictions were then validated against actual sales commission data to assess the model's effectiveness. Throughout the implementation process, we ensured compatibility with various hardware and software requirements, including operating systems, processors, and programming languages/frameworks. Additionally, we developed a user-friendly web interface using technologies such as Flask, HTML, CSS, and Bootstrap, allowing users to easily upload datasets, visualize sales performance, and obtain accurate predictions. Finally, we conducted thorough testing and validation to ensure the reliability and robustness of the system, paving the way for its successful deployment and utilization in real-world sales environments.

5.2 OUTPUT SCREENSHOTS

The 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.

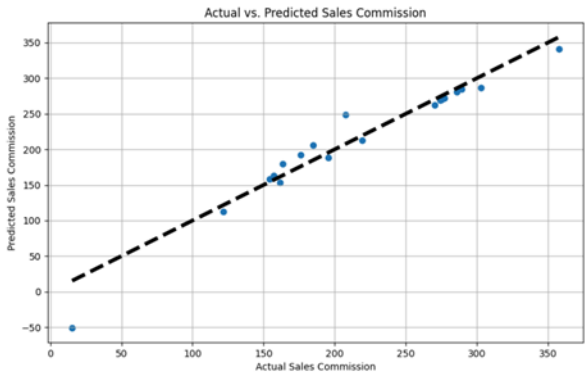

Fig 5.2.1 Actual vs predicted sales commission

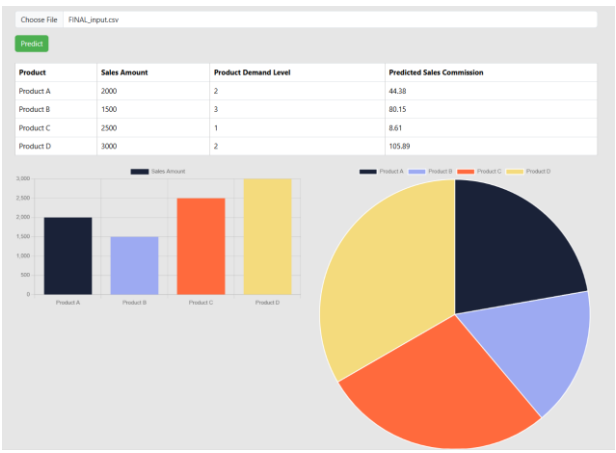

Fig 5.2.2 Output screenshot

CHAPTER 6

CONCLUSION AND FUTURE ENHANCEMENTS

6.1 CONCLUSION

In conclusion, 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.

6.2 FUTURE ENHANCEMENT

This project has a wide range of scope, Dynamic Sales Commission Prediction system could involve incorporating more advanced machine learning algorithms to further improve prediction accuracy and efficiency. Additionally, integrating real-time market data feeds and social media sentiment analysis could provide more comprehensive insights into customer behavior and market trends, allowing for more dynamic adjustment of commission rates. Furthermore, enhancing the user interface and adding features such as personalized dashboards for sales representatives and managers could improve usability and decision-making capabilities. Finally, exploring the possibility of leveraging blockchain technology for transparent and secure commission tracking and payment processing could enhance trust and efficiency in the commission management process.

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] SSongjun 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.