SALES FORECASTING
A PROJECT REPORT

Submitted in the partial fulfilment for the award of the degree of

BACHELOR OF ENGINEERING

IN ARTIFICIAL INTELLIGENCE AND MACHINE LEARNING

Submitted by:

Aryan Gupta
20BCS6656

Ankith Raj
20BCS6684

Aryan Khuswaha
20BCS6691

Sheikh Shahnawaz Hussain
20BCS6628

Under the Supervision of:

Dr.Alankrita Aggarwal

CHANDIGARH UNIVERSITY, GHARUAN, MOHALI - 140413, PUNJAB

10 JANUARY, 2023

CHANDIGARH UNIVERSITY
Discover. Learn. Empower.

BONAFIDE CERTIFICATE

Certified that this project report **" Sales Forecasting "** is the bonafide work of **"** Aryan Gupta , Ankith Raj , Aryan Kushwaha and Sheikh Shahnawaz Hussain**"** who carried out the project work under my/our supervision.

SIGNATURE **SIGNATURE**

HEAD OF THE DEPARTMENT **SUPERVISOR**

Submitted for the project viva-voce examination held on_

INTERNAL EXAMINER **EXTERNAL EXAMINER**

ACKNOWLEDGEMENT

I would like to thank my supervisor, Dr. Alankrita Aggarwal Mam, for her guidance and advice through each stage of making this project and giving me and my teammates Mr. Ankith Raj, Mr. Sheikh Shahnawaz Hussain and Mr. Aryan Khushwaha this opportunity to work on a minor project in which we can show our true potential, creativity, and hard work.

I would also like to thank my family and friends, who have been a constant support and have always motivated me to work hard and bring out the best in me.

This project’s success and end necessitated a great deal of direction and assistance from many people, and we are extremely fortunate to have received it all as part of the project's completion.

We owe everything we've accomplished to their oversight and help, and we'd like to express our gratitude.

ABSTRACT

Forecasting sales is a common and essential use of machine learning (ML). This paper discusses need of Sales Forecasting. Sales Forecasting is the estimate of amount of sales to be expected for a item/product or products for a future period of time. Using Sales Forecasting the management of the enterprise can take decision regarding operations planning, scheduling, production programming inventories of various types, physical distribution and operating profits on the basis of sales forecasts. It also contains some additional benefits like deciding investment proporsals like modernization, expansion of existing units etc. Sales forecasts are essential to make proper arrangement for training the man-power in its own unit or sending them to other industries in the country or abroad to meet the future needs of expertise.

Keywords: sales forecasting, expansion, planning, ML , Training

CHANDIGARH UNIVERSITY
Discover. Learn. Empower.

Table of Contents

List of Tables

Chapter 1: Introduction

Sales forecasting involves analyzing a company's historical sales data to make predictions about its future sales performance. It is an essential aspect of financial planning that helps companies plan for short- and long-term growth. Like all forecasting processes, there is a level of risk and uncertainty involved, so it is important for sales forecasting teams to acknowledge this uncertainty in their forecasts. Sales forecasting is a widely practiced corporate strategy that involves setting objectives, creating action plans, and allocating budgets and resources to achieve those objectives. Several techniques are commonly used for sales forecasting, including Linear Regression, Decision Tree, Random Forest, and XG BOOST.

1.1 PROBLEM DEFINITION

The main focus of this project is to use training data, specifically the data from Big Mart, to createa Sales Forecasting system. By analyzing this data and using machine learning techniques, the goal is to develop accurate models that can predict sales transactions for any company. Sales forecasts can be useful for setting benchmarks, evaluating the impact of new initiatives, and planning resources to meet expected demand. Additionally, they can be used to project future budgets

1.2 PROJECT OVERVIEW

- Sales Forecast helps in predicting the short-term or long-term sales performance of that company.
- The data for the Sales will be taken from Kaggle.
- Data wrangling (Data profiling, missing value treatment, exploratory data analysis) will be performed.
- Prediction will be automated with Machine learning Models thus saving a lot of time.
- We will be working on Pandas, Matplotlib, Model Training, AWS,HTML, and CSS, Flask.

1.3 HARDWARE SPECIFICATIONS

1.3.1 PC

A pc is a personal computer that can be used for multiple purposes depending on its size, capabilities, and price. They are to be operated directly by the end-user. Personalcomputers are single-user systems and are portable. Our web application program willbe installed on the pc for our clients to use it. This makes it feasible for individual use.

1.4 SOFTWARE SPECIFICATIONS

1.4.1 Jupyter Notebook:

Jupyter Notebook is a web-based open-source application that is used for editing, creating running, and sharing documents that contain live codes, visualization, text, and equations. Its core supported programming languages are Julia, R, and Python. Jupyter notebook comes with an IPython kernel that allows the programmer to write programs in python. There are over 100 kernels other than IPython available for use.

1.4.2 Atom Text editor

Atom is a text and source code editor which works across all operating systems. It speeds up find-and-replace operations by an order of magnitude and improves loading performance for large, single-line files It’s a desktop application built with HTML, JavaScript, CSS, and Node.js integration.

1.4.3 AWS

Amazon Web Services, Inc. (AWS) is a subsidiary of Amazon that provides on-demand cloud computing platforms and APIs to individuals, companies, and governments, on a metered pay-as-you-go basis. Through AWS server farms, these cloud computing web services offer software tools and distributed computer processing capability. One of these services is Amazon Elastic Compute Cloud (EC2), which enables customers to have a virtual computer cluster at

their disposal that is always accessible via the Internet. The majority of a real computer's features, such as hardware central processing units (CPUs) and graphics processing units (GPUs) for processing, local/RAM memory, hard-disk/SSD storage, a choice of operating systems, networking, and pre-loaded application software including web servers, databases, and customer relationship management, are all emulated by AWS's virtual computers (CRM).

1.4.4 FLASK

Flask is a micro web framework written in Python. It is classified as a microframework because it does not require particular tools or libraries. It has no database abstraction layer, formvalidation, or any other components where pre-existing third-party libraries provide common functions. However, Flask supports extensions that can add application features as if they were implemented in Flask itself. Extensions exist for object-relational mappers, form validation, upload handling, various open authentication technologies and several common framework- related tools.

1.4.5 MS-EXCEL

Microsoft produced Microsoft Excel, a spreadsheet, for Windows, macOS, Android, and iOS. It has calculating or computing capabilities, graphing tools, pivot tables, and the Visual Basic for Applications macro programming language (VBA). The Microsoft Office programme package includes Excel.

Software Tool Used	Description	Logo
Jupyter Noebook	Jupyter Notebook is a web-based open-source application that is used for editing, creating, running, and sharing documents that contain live codes, visualisations, text, and equations. There are over 100 kernels other than IPython available for use.	
Atom Text Editor	Atom is a text and source code editor which works across all operating systems. It speeds up find-and-replace operations by an order of magnitude and improves performance of files	
Visual Sudio Code	Visual studio code is an open-source code editor built for Windows, Mac OS, Linux which can be used for	

	various programming languages like Java, JavaScript, Python, C, C++, Node.js.	
Flask	Flask is a micro web framework written in Python. It is classified as microframework because it does not require particular tools or libraries. It has no database abstraction layer, formvalidation,or any other components where pre-existing third-party libraries provide common functions.	

2 LITERATURE REVIEW

2.1 Existing System Summary

Year and citation	Annals of R.S.C.B., ISSN:1583-6258, Vol.25, Issue 5, 2021, Pages. 3928 - 3936 Received 15 April 2021; Accepted 05 May 2021.	https://www.anaplan.com/blog/sales-forecasting-guide/ 2022	T uyls, Karl & Maes, Sam & Vanschoenwinkel, B.. (2023). Machine Learning Techniques for sales forecasting
Article Title	"Intelligent Sales Prediction Using Machine Learning Techniques	"MachineLearning Models for Sales Forecasting"	"Forecast of Sales of Walmart Store Using Big Data Applications"
Purpose of the study	The forecast is composed of a smoothed averaged adjusted for a linear trend. Then theforecast is also adjusted for seasonality Machine learning algorithms such as Generalized Linear Model (GLM), Decision Tree (DT) and Gradient Boost Tree (GBT) are used in prediction of future sales.	Regression algorithm captures the patterns in the whole set of stores or products.The analysis includes the attributes such as mean sales value of historical data, state and school holiday flags,distance from store to competitor's store, storeassortment type are considered in prediction.Various machine learning models such as Random Forest, Neural network, Lasso regularization, Arima model and Extra Tree model are used to analyze the data.	The plan calls for gathering vast amounts of sales-related data, which is then sent to HDFS (Hadoop's distributed file system) for map reduction. To forecast sales, the Holt Winters algorithm is employed. The algorithm exhibits seasonality, trend, and randomness. Data sets are utilised to train the algorithm, and then it is used to predict sales.
Tools/ Software used	- Jupyter Notebook	- Jupyter Notebook	- Jupyter Notebook
Comparison of techniques done	- Generalized Linear Model (GLM) - Decision Tree (DT) - Gradient Boost Tree (GBT)	- Lasso - Neural Network	- Neural Network - Decision Tree (DT)
Evaluation parameters	- Model Accuracy	- Model Accuracy	- Model Accuracy

Table 2.1: Literature review summary

2.2 Proposed System

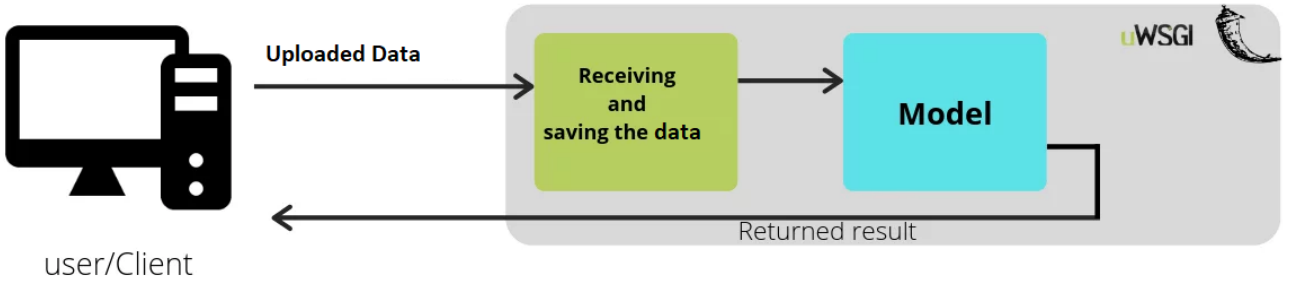
- This project mainly focuses on developing a system that can Predict Sales of a Company
- Data wrangling (Data profiling, missing value treatment, exploratory data analysis) will be performed.
- We will be working on Pandas, Matplotlib, Model Training, AWS,HTML, and CSS.

3 DESIGN PROCESS/FLOW

- Sales forecasting plays a critical role in the success of any business. It provides essential information to allocate resources, hire new staff, manage costs, and increase quotas. The goal of sales forecasting is to provide accurate predictions that businesses can use to make informed and impactful decisions.
- Sales forecasts help businesses make informed decisions about staffing, inventory, product lines, and marketing efforts. It allows sales managers and representatives to spot potential issues and address them before they become problems.
- Sales forecasting is a valuable tool for sales managers and leaders to set realistic goals. Sales forecasts form the basis of your entire strategy throughout the year, and the insights lay the groundwork not just for the company's vision, but also for the direction of sales team.
- Sales forecasting also enables sales managers and leaders to set realistic goals for their teams. It forms the basis of a company's strategy throughout the year and provides insights that guide the direction of the sales team.

4. METHODOLOGY

The following methodology will be followed to achieve the objectives defined for the proposed research work:

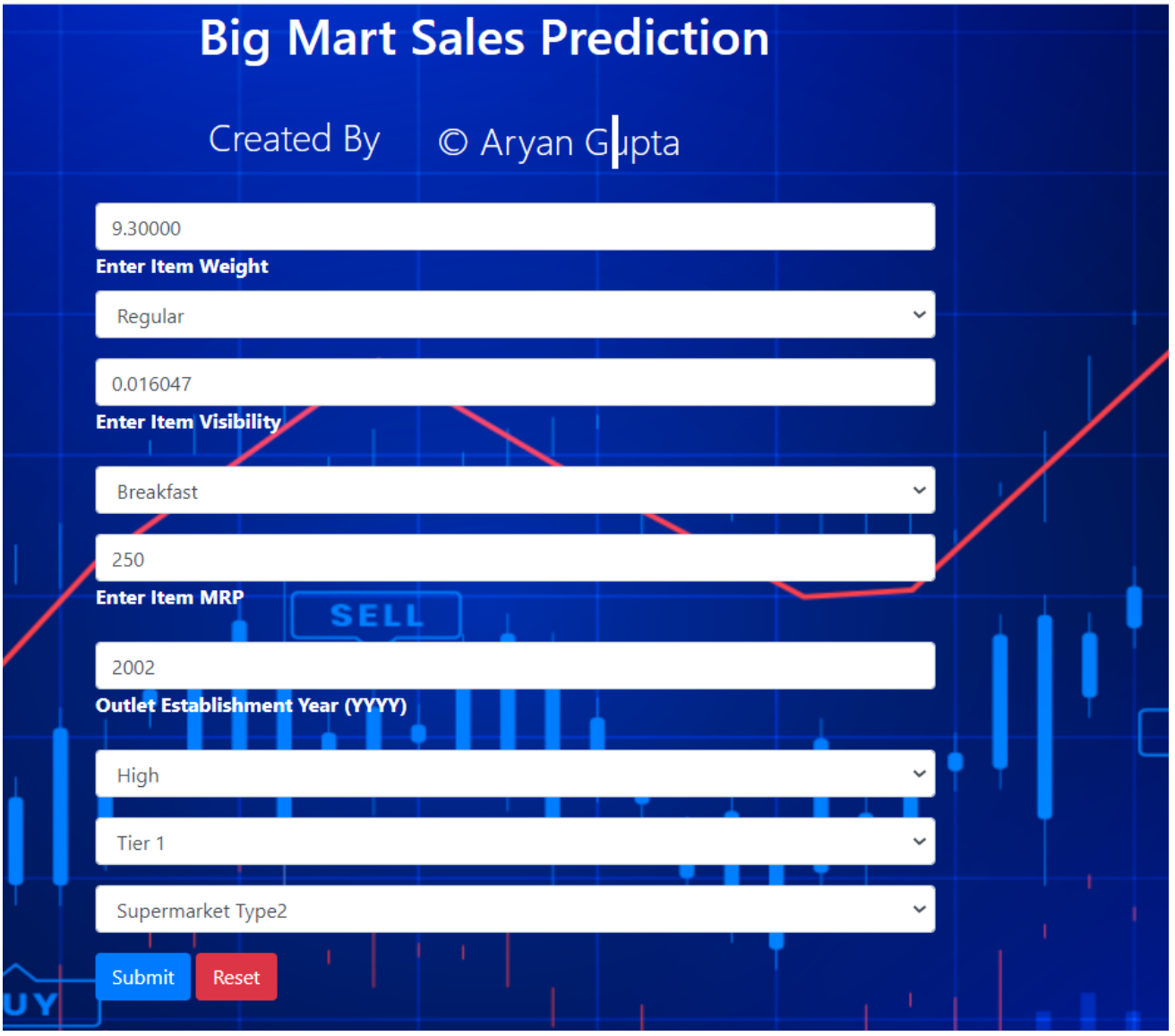


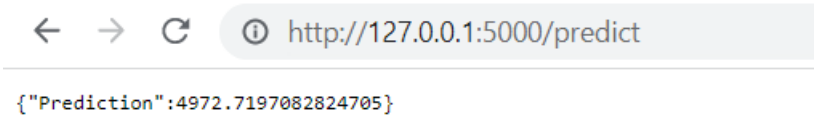
The system flowchart of our project, which is based on Sales Forecasting, is illustrated in the diagram. It is split into two parts: the backend and the frontend. The Bigmart dataset make up the backend element. The Python code is separated into preprocessing, model building, training, and prediction sections.

Flask is a micro web framework written in Python, is what we used for the front-end and connecting to Python Code. Front End is made Using HTML/CSS and consists of various input fields like Item weight, Item visibility, Item Type, Item MRP, Outlet Type, Outlet Location type. User will enter values in these fields and the values will be passed to model and the result of future outlet sales will be shown to the user in JSON Format

5. RESULT ANALYSIS AND VALIDATION

The sales of several Big Mart outlets have been predicted using machine learning methods like Linear Regression, KNearest Neighbors, XGBoost, and Random Forest. For each of the four methods, many metrics that affect the accuracy of results are tabulated, including Root Mean Squared Error (RMSE), Variance Score, Training and Testing Accuracy, and more.





{"Prediction":4972.7197082824705}

6. CONCLUSION AND FUTURE WORK

The usage of machine learning approaches proves to be a crucial feature for designing business plans while taking into consideration consumer buying habits, as traditional methods are not very helpful to commercial organisations in revenue growth. Businesses can adopt effective tactics for expanding sales and stepping unafraid into the competitive world by using sales predictions based on a variety of factors, including past sales. There is a certain degree of predictability when it comes to sales patterns, according to the analysis and modelling done for the sales forecasting project. It is feasible to estimate future sales success by examining previous data and applying the right forecasting tools. However, it's crucial to remember that sales projections are not always accurate and can be impacted by a range of outside factors, including adjustments to the economy, changes in consumer behaviour, and unanticipated events. Overall, sales forecasting is a useful tool for organisations to plan and make defensible judgements about their operations, but it should only be used as a guide. Businesses can more successfully traverse the complexity of the market and accomplish their objectives by fusing the insights received through sales forecasting with other pertinent information and knowledge.

FUTURE WORK

1) Real-time prediction: The ability to predict sales in real-time can be a game-changer for businesses, allowing them to quickly adjust their strategies based on market trends and customer behavior. User can Upload there CSV Files of any organization and get the future sales

2) We intend to expand this research to a larger scale in the future so that different embedding models can be considered on a wider range of datasets

3) To create an android application based on this machine learning model.

4) To integrate this model to an offline based system so that it runs completely without an internet connection

5) To Deploy the Project on Platform like AWS so that anyone with the Link can have access to it.

REFERENCES

[1] https://www.kaggle.com/datasets/shivan118/big-mart-sales-prediction-datasets

[2] https://www.saleshacker.com/sales-forecasting-101/

[3] https://hbr.org/1971/07/how-to-choose-the-right-forecasting-technique

[4] https://www.liveplan.com/blog/the-best-way-to-forecast-sales-and-revenue/

[4] https://repository.upenn.edu/cgi/viewcontent.cgi?article=1186&context=marketing_papers

[5] https://portal.abuad.edu.ng/lecturer/documents/1588001295ICH_254_NOTE_2.pdf

[6] Gudmundsson, K., & Helgason, H. (2020). Machine learning in sales forecasting. Available at: https://www.arnia.com/blog/machine-learning-in-sales-forecasting

[7] Hu, Y., Yang, H., & Chao, C. (2018). Sales Forecasting with a Hybrid Method of ARIMA and XGBoost. In 2018 IEEE International Conference on Systems, Man, and Cybernetics (SMC) (pp. 2813-2818). IEEE.

[8] Sarkar, S., Rana, S., & Kumar, S. (2021). Sales forecasting using machine learning algorithms: A comprehensive review. Journal of Retailing and Consumer Services, 59, 102376.

[9] https://www.yesware.com/blog/sales-forecast/

[10] https://blog.hubspot.com/sales/accurate-sales-forecasting-model-tips

[11] https://whatfix.com/blog/sales-forecasting/