A Mini Project Report on

Sales Prediction

T.E. - I.T Engineering

Submitted By

Jateen Tirlotkar	**20104091**
Jill Shah	**20104004**
Swapnil Sawant	**20104088**

Under The Guidance Of

Ms.Roshna Sangle

DEPARTMENT OF INFORMATION TECHNOLOGY

A.P.SHAH INSTITUTE OF TECHNOLOGY

G.B. Road, Kasarvadavali, Thane (W), Mumbai-400615

UNIVERSITY OF MUMBAI

Academic year : 2022-23

CERTIFICATE

This to certify that the Mini Project report on **Sales Prediction** has been submitted by **Jateen Tirlotkar** (20104091), **_Jill Shah** (20104004) and **_Swapnil Sawant** (20104088) who are a Bonafede students of A. P. Shah Institute of Technology, Thane, Mumbai, as a partial fulfilment of the requirement for the degree in **Information Technology**, during the academic year **2022-2023** in the satisfactory manner as per the curriculum laid down by University of Mumbai.

Ms.Roshna Sangle

Guide

Dr. Kiran Deshpande

Head Department of Information Technology

Dr.Uttam.D.Kolekar

Principal

External Examiner(s)

1.

2.

Place: A.P.Shah Institute of Technology, Thane

Date:

ACKNOWLEDGEMENT

This project would not have come to fruition without the invaluable help of our guide **Ms.Roshna Sangle**. Expressing gratitude towards our HoD, **Dr. Kiran Deshpande**, and the Department of Information Technology for providing us with the opportunity as well as the support required to pursue this project. We would also like to thank our teacher Ms. Charul Singh who gave us her valuable suggestions and ideas when we were in need of them. We would also like to thank our peers for their helpful suggestions.

ABSTRACT

Machine Learning is transforming every walk of life and has become a major contributor in real world scenarios. The revolutionary applications of Machine Learning can be seen in every field including education, healthcare, engineering, sales, entertainment, transport and several more; the list is never ending. The traditional approach of sales and marketing goals no longer help the companies, to cope up with the pace of competitive market, as they are carried out with no insights to customers' purchasing patterns. Major transformations can be seen in the domain of sales and marketing as a result of Machine Learning advancements. Owing to such advancements, various critical aspects such as consumers' purchase patterns, target audience, and predicting sales for the recent years to come can be easily determined, thus helping the sales team in formulating plans for a boost in their business. The aim of this paper is to propose a dimension for predicting the future sales of Big Mart Companies keeping in view the sales of previous years. A comprehensive study of sales prediction is done using Machine Learning models such as Linear Regression, K-Neighbors Regressor, XGBoost Regressor and Random Forest Regressor. The prediction includes data parameters such as item weight, item fat content, item visibility, item type, item MRP, outlet establishment year, outlet size and outlet location type.

TABLE OF CONTENTS

Chapter 1

Introduction

- Sales prediction plays an important role in many fields and helps to improve the sales of a company by making future plans by predicting the sales of a company. Sales prediction is an important prerequisite for enterprise planning and correct decision making, allowing companies to better plan their business activities.In this project, we will use machine learning techniques to analyze the historical sales data of Big Mart and develop a predictive model that can forecast the sales for the future.Previous research on sales prediction has always used a single prediction model. However, no single model can perform the best for all kinds of merchandise.Accurate prediction results for classification in data mining, this research By using the method of decision tree, sellers can boost sales.Several prediction models are included as basic models for classification. We compared the results of the prediction model with other single models.The results show that the accuracy of the prediction model is better than that of a single model. Sales forecasting plays an important role in guiding the sales and marketing of warehousing department planning warehouse location. At an equivalent time, sales data can better reflect future sales trends. Bigmart is a big supermarket chain, with stores all round the country and its current board began a challenge to all or any Data Scientists out there to assist them create a model which will predict the sales, per product, for every store, period of time. The forecasts are generated using the flow of demands from the past as well as by considering other known factors in future. Various machine learning models are developed for the same. The purpose of forecasting sales is mainly to help the organization predict their targets and modify their strategy to improve their productivity in the coming future

1.2 Purpose:

A sales prediction system is a tool that businesses can use to forecast their future sales revenue based on historical sales data, market trends, and other relevant factors. The purpose of such a system is to provide the business with accurate and reliable sales forecasts, which can help them make informed decisions about inventory management, marketing strategies, staffing levels, and other aspects of their operations.

The system works by analysing historical sales data to identify patterns and trends, and then using that information to make predictions about future sales performance. Factors that may be taken into account when making these predictions include past sales volume, seasonal trends, economic indicators, market competition, and customer demographics.

By using a sales prediction system, businesses can better plan for the future, allocate resources more effectively, and make more informed decisions about their sales and marketing strategies. This can ultimately lead to increased revenue and profitability for the business.

1.3 Objectives:

The main objective of a project focused on sales prediction is to use historical data and other relevant factors to forecast future sales. The specific objectives may vary depending on the nature of the project, but some common goals might include:

1. **Accurately predicting future sales figures:** The primary objective of any sales prediction project is to generate accurate forecasts of future sales. This can help businesses plan for demand, allocate resources effectively, and make better-informed decisions about inventory, pricing, and marketing strategies.
2. **Identifying key drivers of sales:** Sales prediction models can help businesses identify the factors that have the biggest impact on sales. By analyzing historical data and other variables, businesses can gain insights into what drives sales and adjust their strategies accordingly.
3. **Optimizing sales strategies:** Once businesses have identified the key drivers of sales, they can use this information to optimize their sales strategies. This might involve adjusting pricing, changing marketing tactics, or making other changes to better align with customer needs and preferences.
4. **Improving forecasting accuracy over time:** Sales prediction models are not static, and businesses should aim to continually refine and improve their models over time. By incorporating new data sources, experimenting with different models, and incorporating feedback from customers and other stakeholders, businesses can improve their forecasting accuracy and make better-informed decisions.

- To determine the amount of product that will be required in future.
- To compare and evaluate the performance of prediction algorithms.
- To create a machine learning model that can help Big Mart improve their business operations and stay ahead of the competition
- To develop an accurate and reliable model that can predict sales for each product in each store.

1.4 Scope: -

The scope for Big Mart sales prediction using machine learning is vast and significant. By applying machine learning techniques to sales data, it is possible to accurately predict future sales, identify trends and patterns, and optimize sales strategies. This can help retailers make informed decisions about inventory management, pricing, and marketing, which can ultimately lead to increased revenue and profitability. Some specific areas of application for Big Mart sales prediction using machine learning include:

Sales forecasting: Machine learning models can be trained to analyze historical sales data and predict future sales trends, allowing retailers to adjust their inventory and staffing accordingly.

Customer segmentation: Machine learning algorithms can be used to group customers based on their buying behavior, demographics, and preferences, allowing retailers to personalize their marketing efforts and optimize their sales strategies.

Product recommendations: Machine learning models can be used to analyze customer purchase history and recommend products that are likely to be of interest to them.

Price optimization: Machine learning can be used to identify optimal pricing strategies based on customer behavior, competitor pricing, and market trends.

Overall, the application of machine learning to Big Mart sales prediction has the potential to revolutionize the retail industry, improving efficiency and profitability for retailers and enhancing the shopping experience for customers.

Chapter 2

Literature Review: -

Title	Author	Year	Algorithm's Used	Result
BigMart Sales Prediction	Akshay Arora, Sudhanshu Kumar	2021	Random Forest, Linear Regression, Decision Tree.	Random Forest achieved highest accuracy.
Sales Prediction of BigMart using Machine Learning	Manisha Borkar, Kalayani Thakre	2020	K-nearest neighbor, Support vector regression.	Support vector regression achieved best accuracy.
BigMart Sales using Machine Learning techniques.	Mitali Awari, Dr. Dlip Palange.	2018	Random Forest, XgBoost, Ridge regression.	Random Forest achieved best accuracy.

Chapter 3

Proposed System: -

A sales prediction system is a software application or machine learning model that uses historical sales data and other relevant variables to generate forecasts of future sales. The proposed system of sales prediction might include the following components:

1.Data collection: Collecting historical sales data from Big Mart stores, including information on product characteristics, store location, and sales figures.

2.Data preprocessing: Cleaning and transforming the data to ensure it is in a suitable format for analysis. This may involve removing missing values, dealing with outliers, and encoding categorical variables.

3.Feature selection: Identifying the most relevant features for sales prediction, based on their correlation with sales and their contribution to the model's accuracy.

4.Model development: Training a machine learning model using the selected features and historical sales data. The choice of model would depend on the specific business needs and data available, but could include linear regression, decision tree, random forest, or neural network.

5.Model evaluation: Evaluating the performance of the model using metrics such as mean absolute error, mean squared error, and R-squared. The model may need to be fine-tuned or retrained based on the evaluation results.

6.Sales prediction: Using the trained model to predict future sales figures for each product in each store, based on the relevant features.

3.1 Features and Functionality: -

Sales prediction is the process of using data and analytical techniques to forecast future sales of a business. The features and functionality of sales prediction can vary depending on the specific software or tool used, but some common ones include:

Data integration: Sales prediction software should be able to integrate data from various sources such as customer data, inventory data, and financial data, to provide a comprehensive view of the business.

Historical data analysis: The software should analyze historical sales data to identify patterns and trends, which can be used to make accurate predictions about future sales.

Machine learning algorithms: Many sales prediction tools use machine learning algorithms to analyze data and make accurate predictions.

Forecasting: Sales prediction software should provide accurate sales forecasts based on historical data, market trends, and other factors.

Data visualization: The software should be able to present data in an easy-to-understand format, such as graphs and charts, to help businesses make informed decisions.

What-if scenarios: Sales prediction tools should allow businesses to test different scenarios and see how they would impact sales. For example, they might test how a price increase or a new marketing campaign could affect sales.

Customization: Some sales prediction tools offer the ability to customize the model used to make predictions, allowing businesses to fine-tune the tool to their specific needs.

Overall, sales prediction tools help businesses make informed decisions about inventory management, marketing, and other critical areas by providing accurate sales forecasts

Chapter 4

Project Scheduling Template

Sr. No	Group Member	Time duration	Work to be done
1	Jateen Tirlotkar	1st week of January	Implementing 1st module/ functionality (Data Collection Module)
		2nd week of January	Testing 1st module (Data Cleaning and Preprocessing Module)
2	Jill Shah	3rd week of January	Implementing 2nd module/ functionality (Sales Prediction Model Evaluation Module, Sales Forecasting Module)
3	Swapnil Sawant	By the end of march month	Implementing 3rd module/ functionality (Reporting Module, Integration Module)

Chapter 5

Requirement analysis: -

Understand the Business Problem: The first step is to understand the business problem that needs to be solved. In this case, the business problem is to predict the sales for a given product or service based on historical data.

Gather Data: The next step is to gather relevant data that can be used to train the ML model. This may include data on past sales, product or service features, customer demographics, and other relevant factors that may affect sales.

Preprocess the Data: Once the data has been gathered, it needs to be cleaned, transformed, and preprocessed to make it suitable for training the ML model. This may include removing missing values, transforming categorical variables into numerical features, scaling the data, and other steps.

Select the ML Model: There are many different ML models that can be used for sales prediction, including linear regression, decision trees, neural networks, and others. The choice of model will depend on the specific problem and the nature of the data.

Train the ML Model: Once the model has been selected, it needs to be trained using the preprocessed data. This involves selecting appropriate training algorithms, tuning hyperparameters, and evaluating the performance of the model.

Validate the Model: After training, the model needs to be validated to ensure that it is accurate and robust. This involves testing the model on a separate dataset, and comparing the predicted sales to the actual sales.

Deploy the Model: Once the model has been validated, it can be deployed in a production environment. This may involve integrating the model into an existing system, or building a new system to handle sales predictions.

Monitor the Model: Finally, it is important to monitor the performance of the model over time, and update it as necessary to ensure that it continues to provide accurate predictions.

In summary, a sales prediction project using ML involves understanding the business problem, gathering and preprocessing data, selecting and training an appropriate ML model, validating the model, deploying it in a production environment, and monitoring its performance over time.

Chapter 6

Project Design

6.1 Use Case diagram:-

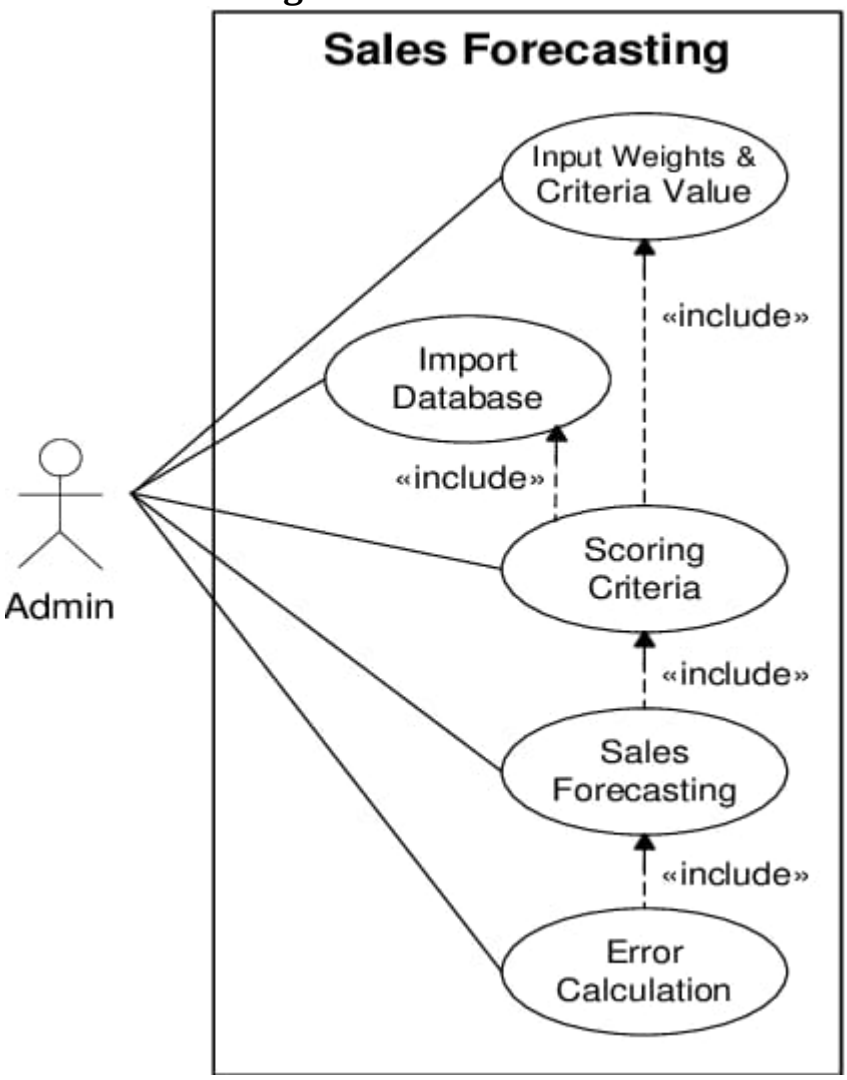


Fig .1 User case diagram

1. Predict Sales: This is the main use case for the system, which involves predicting future sales projections based on historical sales data.

2. Data Pre-Processing: This use case involves preparing the input data for the sales predictor system by cleaning, filtering, transforming, and normalizing the data.

3. Sales Data Input: This use case involves entering, importing, verifying, and storing sales data in the system.

4. Sales Data Output: This use case involves viewing and exporting the sales data from the system.

5. Sales Reports: This use case involves generating, viewing, and saving reports based on the predicted sales data.

6.2 Data Flow Diagram (DFD):-

- **Level 0 DFD**

Fig.1 DFD Level 0

- **Level 1 DFD**

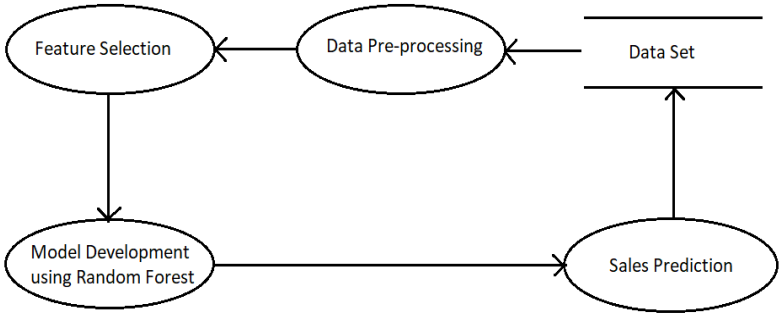


Fig.2 Level 1 DFD

6.3 System Architecture: -

A system architecture for a sales prediction project typically involves several components that work together to process and analyze data, make predictions, and present results to users. Here is a possible architecture:

Data Collection: The first step is to collect sales data from various sources such as point of sale systems, customer relationship management systems, and inventory management systems. The data can be collected either in real-time or batch mode.

Data Preprocessing: The raw data is then preprocessed to clean and transform it into a format that can be easily processed. This step involves handling missing data, normalizing data, and removing outliers.

Feature Engineering: The preprocessed data is then transformed into meaningful features that can be used for analysis. These features may include customer demographics, sales history, seasonality, promotions, and other factors that can influence sales.

Modeling: The next step is to build a predictive model using machine learning algorithms such as regression, neural networks, or decision trees. The model is trained on historical data to learn patterns and make predictions.

Evaluation: The model is then evaluated using a testing dataset to assess its accuracy and performance. The evaluation metrics can include Mean Absolute Error (MAE), Mean Squared Error (MSE), or Root Mean Squared Error (RMSE).

Deployment: The trained model is then deployed in a production environment where it can make predictions on new data. This step involves integrating the model into an application or web service.

Monitoring: Finally, the model performance is continuously monitored to ensure that it is producing accurate and reliable predictions. This step may involve retraining the model periodically or updating the features used for prediction.

The architecture described above is a high-level overview of the components involved in a sales prediction project. The actual implementation of the architecture will depend on the specific requirements and constraints of the project.

Chapter 7

Technical Specification: -

Development: VS Code

VS Code also known as Visual Studio Code is a source code editor made by Microsoft for Windows, Linux, MacOS. It has various features such as Debugging, Syntax highlighting, extension, intelligent code completion.

Frontend: Html, CSS, JavaScript

As a web developer, the three main languages we use to build websites are HTML, CSS, and JavaScript. JavaScript is the programming language, we use HTML to structure the site, and we use

CSS to design and layout the web page. These days, CSS has become more than just a design language, though. You can actually implement animations and smooth transitions with just CSS.

OS: Windows

Windows is a **graphical operating system** developed by Microsoft. It allows users to view and

store files, run the software, play games, watch videos, and provides a way to connect to the internet.

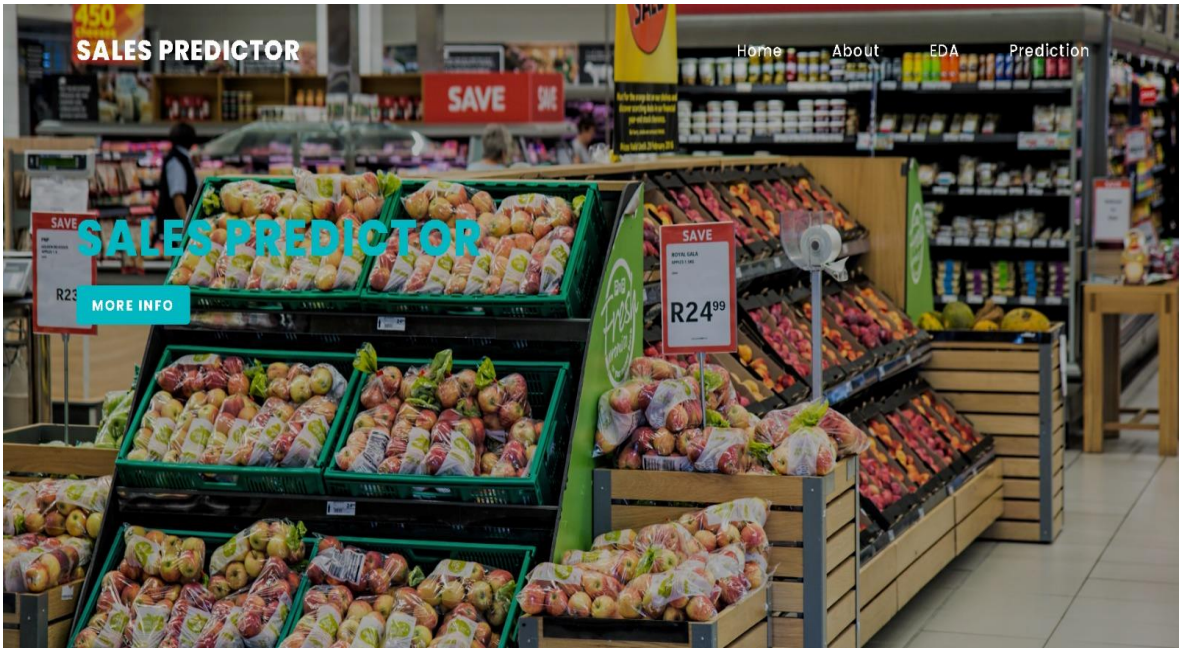
It was released for both home computing and professional works.

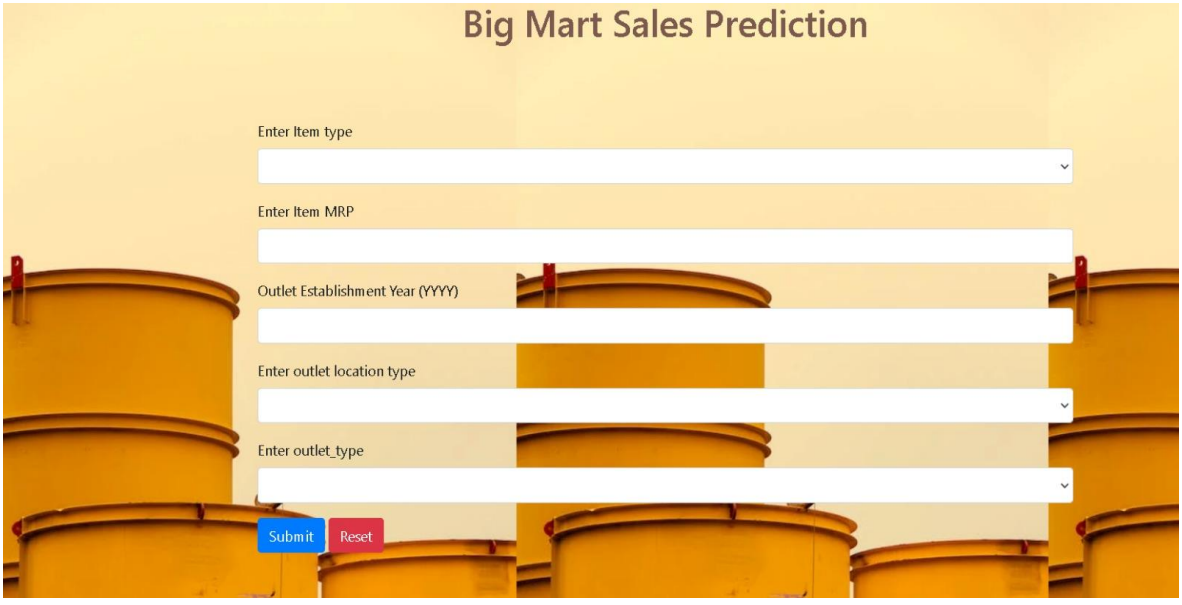
Backend:

Python Using ML Algorithms Using Python for developing a hate speech detection and fake news detection system is a common choice as Python provides many powerful libraries and tools for natural language processing and machine learning. Here are some of the commonly used libraries and tools for building such a system: LogisticRegression DecisionTreeClassifier GradientBoostingClassifier RandomForestClassifier SnowballStemme

Chapter 8

Implementation:





About

Sales forecasting software uses quantitative methods to analyze historical business data and trends – such as closed and won deals and win/loss records – and then produces an accurate report of expected sales revenue. Forecast reports compare sales targets with achieved sales versus expected sales. Businesses need accurate sales forecasts to help allocate resources, hire new staff, increase quotas, and manage costs. Sales forecasting tools enable companies to predict future growth trends and help leadership formulate effective strategies to expand their business. Sales forecasting software helps businesses answer these questions: What is our expected revenue? (Organized by salesperson, territory, or account.) How did actual sales compare to expected sales? Which method produces the most accurate forecast? Furthermore, publicly-traded companies have to report an accurate sales forecast because they know investors are looking for success, reliability, and stability each quarter. The sales forecasting software models are well-tested and robust, but users can adjust the models to tweak predictions and make them more accurate over time. Businesses can also modify the forecasts with "what if assumptions" to simulate sales behavior and market conditions.

EDA

Exploratory Data Analysis

Value	Count	Frequency (%)
Fruits and Vegetables	1232	14.5%
Snack Foods	1200	14.1%
Household	910	10.7%
Frozen Foods	856	10.0%
Dairy	682	8.0%
Canned	649	7.6%
Baking Goods	648	7.6%
Health and Hygiene	520	6.1%
Soft Drinks	445	5.2%
Meat	425	5.0%

Chapter 9

Conclusion & Future Scope:

In conclusion, a sales prediction system can provide businesses with valuable insights into their sales performance, enabling them to make informed decisions about inventory management, marketing strategies, and other aspects of their operations. By analyzing historical sales data and incorporating machine learning algorithms, these systems can generate accurate and reliable sales forecasts that help businesses plan for the future and improve their profitability.

Looking ahead, there is significant potential for further advancements in sales prediction systems. Some possible future developments include:

1. Integration with AI and IoT: Sales prediction systems can be enhanced by integrating with artificial intelligence (AI) and Internet of Things (IoT) technologies to enable real-time data analysis and predictive modeling.
2. Personalized recommendations: Sales prediction systems can be leveraged to provide personalized recommendations to customers based on their purchase history and preferences, improving customer satisfaction and loyalty.
3. Enhanced data visualization: Further advancements in data visualization techniques can improve the ability of sales prediction systems to provide actionable insights and help users identify trends and patterns.
4. Expansion into new markets: Sales prediction systems can be expanded into new markets and industries, such as healthcare and finance, where they can help organizations optimize their operations and improve their bottom line.

Overall, the future of sales prediction systems is bright, with new advancements and applications continuing to emerge. As businesses increasingly rely on data-driven insights to make informed decisions, sales prediction systems will play an increasingly important role in driving their success.

Chapter 10

References

[1] Abror Abduvaliyev, Al-Sakib Khan Pathan, Jianying Zhou, Rodrigo Roman and WaiChoong Wong ,“On the vital Areas of Intrusion Detection Systems in Wireless Sensor networks”,IEEE Communications Surveys & Tutorials, Accepted For Publications, 2013-in press.

[2] H.H. Soliman, et al,“A comparative performance evaluation of intrusion detection techniques for hierarchical wireless sensor networks”, Egyptian Informatics Journal (2012) 13, 225238.

[3] Giannetsos Athanasios, “Intrusion Detection in Wireless Sensor Networks”, Master THESIS, Carnegie Mellon University, April 8, 2008.

[4] K.Fall and K.Varadhan,“The NS Manual”, http://www.isi.edu/nsnam/ns/doc/ns doc.pdf., 1 Feb 2014.

[5] Jae Chung and Mark Claypool,“NS by Example-Tutorial”, http://nile.wpi.edu/NS/overview.html , 1 Feb 2014.

[6] Network Simulator blog, http://Mohittahilani.blogspot.com , 1 Feb 2014. [7] AWK Script for NS2, http://mohit.ueuo.com/AWK-Scripts.html , 1 Feb 2014.