INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

Impact Factor: 8.165

Future Sales Prediction using Deep Learning

Archita Gujare, Shubham Bhattad, Sejal Rajadhyaksha, Ishwari Dhotre, Prof. N.R. Sonawane

Department of Information Technology, PVG's COET, Pune, India

Department of Information Technology, PVG's COET, Pune, India

Department of Information Technology, PVG's COET, Pune, India

Department of Information Technology, PVG's COET, Pune, India

ABSTRACT: Integration of decision analysis and forecasts is required by an Intelligent Decision Analytical System. The majority of businesses rely largely on a knowledgebase and demand forecasting of sales patterns. The accuracy of a sales estimate has a significant influence on a company's bottom line. Data mining and machine learning methods are powerful tools for extracting hidden information from massive amounts of historical data in order to improve predicting accuracy and efficiency. This study examines and analyses intelligible predictive models in order to enhance sales forecasting in the future. Traditional forecasting systems struggle to cope with massive data and sales forecasting accuracy. Various data mining approaches might be used to solve these problems.

KEYWORDS: LSTM, Big Mart, Futures sales prediction, Machine Learning

I. INTRODUCTION

1.1 Background

Every company's lifeblood is sales, and sales forecasting is an essential part of running a firm. By expanding market knowledge, good forecasting aids in the development and improvement of corporate strategy. A typical sales forecast delves carefully into prior scenarios or conditions, draws inferences about client acquisition, and analyses inadequacies and strengths before deciding on a budget and marketing strategy for the future year.

In other terms, sales forecasting is the projection of future sales based on previously known resources. An in-depth understanding of previous resources enables the organisation to plan for future demands and raises the chance of success regardless of external factors. Businesses that prioritise sales forecasting are more likely to succeed than those that do not.

1.2 Problem statement of the project:

Integration of decision analysis and forecasts is required by an Intelligent Decision Analytical System.Many of the business organizations are mostly depend on a knowledge base and demand prediction of sales trends. The accuracy of a sales estimate has a significant influence on a company's bottom line.

1.3 Objectives of the Project

1] Understanding of key factors in sales forecasting

2] Design of Sales Prediction model

3] Analysis of the metric used for evaluation.

II. LITERATURE SURVEY

1. MohitGurnani et al. [1] suggested using a variety of machine learning models, hybrid models, and deconstruction techniques to anticipate Rossmann Store sales. Our dataset was first subjected to a linear model, such as ARIMA. Other linear models are excluded since they do not enable external regressors to be included. The results revealed that ARIMA was unable to capture nonlinear patterns correctly. As a result, nonlinear models like as ARNN, SVM, and XGBoost are utilised to address this flaw.

2. Akshay Krishna et al. [2], attempting to estimate retail shop sales using several machine learning approaches and determining the optimum solution for our specific issue statement. In our approach, we used both traditional regression and boosting approaches, and we discovered that the boosting algorithms outperformed the traditional regression algorithms.

3. Gopalakrishnan T et al. [3], analyse a large superstore's sales and forecast future sales in order to assist them boost profitability and improve their brand's competitiveness in light of market changes while also increasing customer happiness.

4. SamanehBeheshti-Kashi et al. [4] offer state-of-the-art sales forecasting methodologies with an emphasis on fashion and new product forecasting. This research also looks at several approaches to predicting the predictive value of user-generated content and search queries.

5. Sandeep Chavan et al. [5] suggested a machine learning-based prediction model for effectively forecasting online goods sales. Our research tries to anticipate product sales using up-to-date data such as online reviews, online ratings, online advertising techniques and feelings, and a variety of other characteristics.

6. ErkinGule et al. [6] uses data mining tools to solve difficulties and approaches for agricultural machinery and machine selling prices. The effective finding of previously discovered patterns in massive datasets is known as data mining. Data gathering, preprocessing, data exploration and model development, interpretation, and assessment are all part of this dynamic information discovery process.

7. Norulhidayah Isa et al. [7] described all of the actions and procedures carried out by the researcher in order to meet all of the project's goals.

8. HerambKadam et al. [8], based on historical sales volume data, provide a projection for large mart sales data in a variety of big mart retailers across different site types. To anticipate sales volume, we may apply the methods of multiple linear regression analysis and random forest, depending on the peculiarities of the data.

III. ALGORITHM

1. LSTM

Long short-term memory (LSTM) is a form of RNN in which the layers' memory is greater, making it more suited to longer-term predictions. The model is made up of cells where data is stored, processed, and then moved on to the next cell. Each cell has many gates that decide whether data is saved, erased, or sent to the next cell.
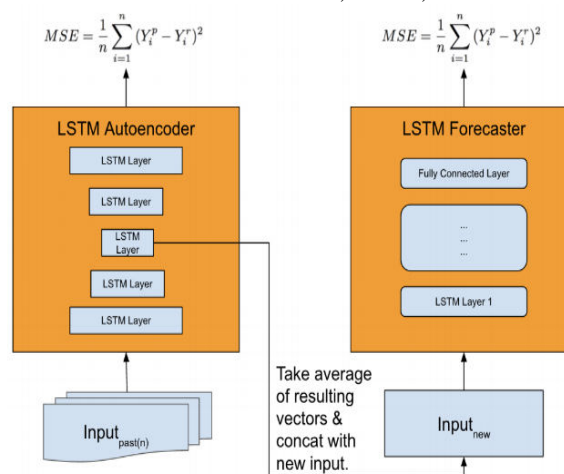

Figure 1. LSTM

This algorithm work tremendously well on a large variety of sequence modelling problems, and they are widely used. LSTM ismainly designed to avoid the long-term dependency problem. Remembering information for long periods of time is their default behavior. Lets see how an RNN looks As we saw in the RNN article the RNN unit takes the current input (X) as well as the previous input (A) to produce output (H) and current state (A).
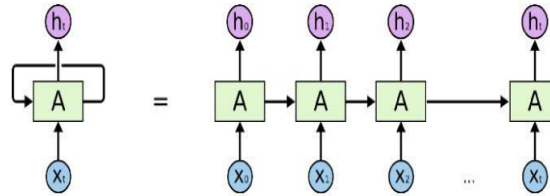
LSTMs also have a similar structure like RNN though the internals have different components as compared to a single tanh (activation) layer in the RNN. There are 4 layers inside an LSTM block which interact together
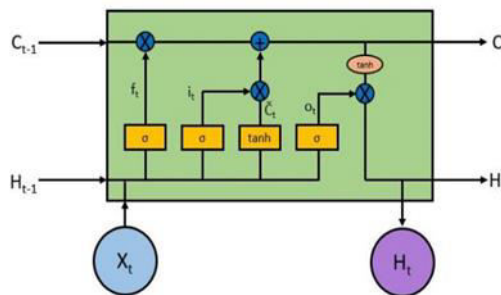
At first it looks pretty complicated and intimidating but lets break it down and understand what is the purpose of each layer and block. The key operation of LSTM is the top horizontal line running from left to right enclosed in the highlight below. With some minor linear interactions along the following line the cell state C enables information to run through the entire LSTM unchanged which enables LSTM to remember context several time steps used before. This line consists of several inputs and outputs which allow us to add or remove information to the cell state. The gates control the removal or addition of information. These layers are called the sigmoid layers (Yellow boxes inside the RNN cell). Their outputs are numbers between zero and one, describing how much of each component should be let through. A value of zero means let nothing through, while a value of one means let everything through. An LSTM consists of three of these gates which are in charge to control the cell state.

IV. PROPOSED SYSTEM

The flow diagram in fig. 1 depicts the phases that the Big Mart sales dataset goes through in order to build up the suggested model and deliver correct results. There are seven phases in all, and each one is important in constructing the suggested model, which is a two-level statistical model.
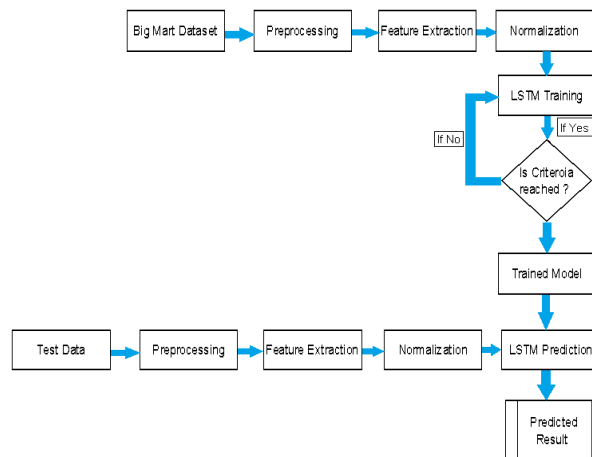

Figure 2. System Architecture

The first five phases are a pre-processing phase before constructing the single model as well as the stack model, and when the model is built, it is evaluated with unseen data to determine its correctness. The lower the mean absolute error, the better the model.

A conventional sales forecast delves carefully into prior events or conditions, draws inferences about client acquisition, and analyses inadequacies and strengths before establishing a budget and marketing strategy for the future year. In other terms, sales forecasting is the projection of future sales based on previously known resources. An in-depth understanding of previous resources enables the organisation to plan for future demands and raises the chance of success regardless of external factors. The user must first register and log in to the system using the appropriate credentials. Users operate on the BigMart dataset product, while the system performs tasks like as data pre-processing, feature selection, normalisation, check criteria, and LSTM prediction. After you've completed all of the tasks, look at the projected outcome. In both the training and testing modules, log out of the system. The BigMart sales information is used to create the suggested model, which produces accurate results. There are seven phases in all, and each one is important in constructing the suggested model, which is a two-level statistical model. The first five phases are a pre-processing phase before constructing the single model and stack model, then once the model is built, the model is tested with unseen data to determine its correctness. The better the model, the lower the value of the mean absolute error.

V. DATASET

We have used a super store dataset for future sales prediction which which consisting of sales and store information from 2013 to 2016 including 3000 rows and around 24 columns.

VI. CONCLUSION

Every business wants to know what its customers want in advance of any season to prevent product shortages. The requirement for organisations to be more precise in their forecasts will grow tremendously as time goes on. As a result, much study is being conducted in this field in order to produce reliable sales estimates. The profit earned by the firm is directly proportionate to the accuracy of the projections. In this study, an attempt was made to properly estimate product sales from a particular outlet using a machine learning model that minimizes the mean absolute error value.

REFERENCES

[1] MohitGurnani , YogeshKorkey , PrachiShahz , Sandeep Udmalex , Vijay Sambhe{ , and Sunil Bhirudk, "Forecasting of sales by using fusion of Machine Learning techniques ",2017 International Conference on Data Management, Analytics and Innovation (ICDMAI) Zeal Education Society, Pune, India, Feb 24-26, 2017.

[2] Akshay Krishna, Akhilesh V, AnimikhAich, ChetanaHegde , "Sales-forecasting of Retail Stores using Machine Learning Techniques", 3rd IEEE International Conference on Computational Systems and Information Technology for Sustainable Solutions 2018

[3] Gopalakrishnan T "Prediction of Sales Value in online shopping using Linear Regression", 2018 4th International Conference on Computing Communication and Automation (ICCCA)

[4] SamanehBeheshti-Kashia,b∗, Hamid Reza Karimic , Klaus-Dieter Thobenb, Michael Lütjenb and Michael Teuckeb, "A survey on retail sales forecasting and prediction in fashion markets", Systems Science & Control Engineering: An Open Access Journal, 2015 Vol. 3, 154–161

[5] Sandeep Chavan, "Predicting Online Product Sales using Machine Learning", International Journal of Engineering Research & Technology (IJERT) http://www.ijert.org ISSN: 2278-0181 Vol. 9 Issue 04, April-2020

[6] ErkinGuler, 2 TanerErsoz and 1 FilizErsoz, "Applying Data Mining Technique to Sales Forecast",

[7] Norulhidayah Isa1 , NurSyuhadaMohd Yusof2 , Muhammad Atif Ramlan3 , "The Implementation of Data Mining Techniques for Sales Analysis using Daily Sales Data ", International Journal of Advanced Trends in Computer Science and Engineering, Volume 8, No.1.5, 2019

[8] 1Heramb Kadam, 2Rahul Shevade, 3 Prof. DevenKetkar, 4Mr. SufiyanRajguru, "A Forecast for Big Mart Sales Based on Random Forests and Multiple Linear Regression", IJEDR 2018 | Volume 6, Issue 4 | ISSN: 2321-9939

[9] 1 Inedi. Theresa, 2Dr.Venkata Reddy Medikonda, 3K.V. Narasimha Reddy, " Prediction of Big Mart Sales Using Exploratory Machine Learning Techniques", International Journal of Advanced Science and Technology Vol. 29, No. 6, (2020), pp. 2906 – 2911

[10] Bohdan M. Pavlyshenko , "Machine-Learning Models for Sales Time Series Forecasting"

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

📱 9940 572 462 🟢 6381 907 438 ✉ ijircce@gmail.com