Predicting Online Marketing Performance with Machine Learning Models

K.C.Bhanu[1], Dr.P.Uma Maheswari Devi[2]
[1]Research Scholar, Department of Commerce and Management Studies
Adikavi Nannayya University ,Rajahmundry
[2]Associate Professor , Department of Commerce and Management Studies
Adikavi Nannayya University ,Rajahmundry
bhanu1605@gmail.com,umadevi_4@yahoo.com

ABSTRACT

Now-a-days Online shopping is the method through which customers deal directly with products, services, etc. via the internet, directly from a seller in real-time interaction without the need of a middleman. E-commerce, or electronic commerce, is what is done when there is an intermediary. The primary goal of this project is to estimate the sales of various products, including groceries, electronics, autos, and more, using a machine learning algorithm. Sales forecasting is another name for projecting sales. In the process of projecting how much product or service a single salesperson, a sales team, or a firm is expected to sell in a specific time period, such as a day, week, month, quarter, or year, future sales are estimated. Today's market benefits both the manufacturer and the several other enterprises that produce components from sales forecasts. Additionally, forecast similar product recommendations by utilising CNN based on customer evaluations of things with visually comparable features, such as those you find on websites, such as Product that is like this.

Keywords: online shopping, machine learning, sales, predict, companies, similar products, product recommendation

INTRODUCTION

The study of computational and statistical methods that automate the process of knowledge acquisition through experience is known as machine learning. There have been many different kinds of investigations into sales forecasting, some of which have been broken down and explored. The process of obtaining sales data from a store over a period of time and making projections is known as sales forecasting. The use of several prediction techniques. There are several factors that affect sales forecasting, some of which are listed here. Country and town vacations, population fluctuations, sales campaigns, and so on are all examples of direct and indirect rivalry.

The aforementioned elements lead to a significant variation in the system's ability to estimate revenue, making it less accurate than expected. All algorithms no longer meet the standard of self-belief. The importance of the Christmas season to commerce. No thought is given to prediction. As a result, the sales vary depending on the system learning methods used.

Over the last ten years, internet commerce has evolved significantly, and more sectors are now specialising in it than ever before. However, opening an online business and beginning to sell right away may not be successful. To understand client preferences and determine what would be fashionable for company, several machine learning and data mining techniques are required. This paper builds an income prediction version appropriate for online products and makes a speciality of evaluating the model's adaptability in specific types of online products. It does so in accordance with the decision-making requirements of online product income, as well as the influencing factors of online product income in various industries and the advantages of deep learning algorithms. Currently, there is a wide variety of research approaches and reasonably extensive research outcomes on product sales forecasting. A sustainable proportion of revenues are generated by major online shopping holidays like Diwali, Christmas, New Year, and Pongal for an online company. The firm employed a large number of temporary workers as well as more permanent personnel in order to better serve customers throughout the Christmas shopping season. This Paper discuss is to forecast internet sales using machine learning. The data contains significant patterns and variables that may be modelled using machine learning algorithms, which can estimate sales with a high degree of accuracy. An effective forecasting model may significantly boost supermarket profits and is often very important to the association since it increases profits and gives insight into how consumers can be better serviced. Mechanism for recommending related products is also predicted. A recommender system, also known as a recommendation system, is a kind of information filtering system that suggests content that will be most useful to a certain user. The recommendations often relate to interesting decision-making processes, such as what to purchase, what music to listen to, or what online news to read. When a person has to choose an item from a service's potentially tempting array of options, recommendation algorithms are very helpful.

Past data is gathered and analysed using quantitative or qualitative models for sales forecasting in order to analyse, identify, and forecast trends that will help with direct demand planning, future operations, and marketing operations. employing a machine learning algorithm to anticipate sales and determine whether they will rise or fall. The major goal of the recommendation system is to anticipate client attraction and propose products that are somewhat comparable to that customer's interest. Since CNNs do not need human supervision to detect key traits, we here propose the comparable product that uses CNN. CNNs are excellent at both classification and image recognition. The weight sharing feature of CNNs is another benefit. In compared to a standard neural network, it also minimises the computation. The primary benefit of this approach is that it may rapidly increase the similarity between fresh material that is seen by consumers who have similar interests. *The problem of finding similar data while looking the internet represents a big challenge for web users due to the enormous quantities of available information on the web[* Noaman M. Ali; Abdullah Alshahrani *]*
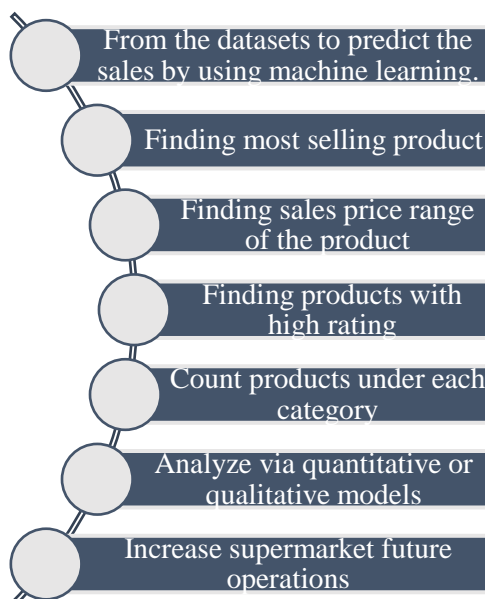
The machine learning algorithm is used to analyse alternative methods for predicting sales using datasets of daily shop sales in order to fix the flaws in the current system. By predicting store sales using well-trained datasets, the system expands its functionality. Where the store uploads this deals map, the result is based on the best algorithms that deliver the best results

with high accuracy. A pie chart and graph are used to better visualise all the precision. The use of machine learning and linear regression increases the system's effectiveness and makes it more equipped to manage large datasets for the store. Additionally, it teaches the shopkeeper how to enhance his promotions and meet client requests. Likewise forecast a similar product suggestion system. A recommender system, also known as a recommendation system, is a kind of information filtering system that makes recommendations for details most relevant to a certain user.Time series data on a firm are required in order to estimate future sales of that company or its product.

Sales prediction is an important part of different companies engaged in merchandising logistics, manufacturing,marketing and merchandising. It allows companies to efficiently allocate coffers, to estimate attainable deals profit and to plan a better strategy for unborn growth of the company[Punam, K:Pamula].

An essential factor in product manufacture is sales forecasting.
- In this study, we develop sales predictions that will aid companies in making wiser choices in the future. The ability to predict things like profit or loss, accuracy, training, testing, validation, which product will generate the most sales, and discounts within a certain time frame.
- An effective forecasting model is often very important to the organisation as it enhances profit and gives insights into how consumers may be better serviced. It can significantly boost supermarket sales. Machine learning is a tool for sales forecasting.
- Due to the vast amounts of information accessible on the online, people have a significant hurdle when trying to discover comparable facts when browsing the internet. Based on consumer evaluations of visually like products, such as those you see on websites, like this one, we forecast similar product recommendations.

From the datasets to predict the sales by using machine learning.

Finding most selling product

Finding sales price range of the product

Finding products with high rating

Count products under each category

Analyze via quantitative or qualitative models

Increase supermarket future operations

DATASET

Collected the internet-based datasets for the kaggle.com webpages. In this study, there are test datasets and train datasets with a total of 5000 test datasets and 8000 train datasets.

Attributing missing numbers and handling outliers are often steps in this process. In our situation, we take the mode and mean of the appropriate characteristic for each missing value, taking into account their kind, which reduces the correlation between input attributes. Advanced tree-based algorithms are resistant to outliers, despite the fact that outliers are unquestionably significant in machine learning methods. The insinuation method will thereafter be our main emphasis since it is a crucial step in attributing missing values in the data and looking for outliers.

METHODS
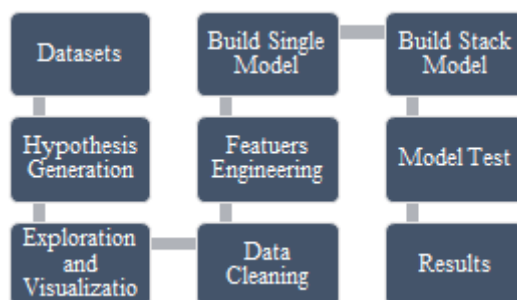
MACHINE LEARNING

The AI system does its duty by using a machine learning algorithm, which typically forecasts output values from the provided input data. Classification and regression are the two basic techniques used by machine learning systems. It enables companies to create more complex forecasting models that make use of more datasets with the least amount of manual labour. By using machine learning algorithms on their data, businesses may enhance the way their goods and services address customer needs.

SUPERVISED MACHINE LEARNING

Supervised learning is a subset of machine learning in which the output is predicted by the machines after they have been trained on well-labeled training data. The term "labelled data" refers to input data that has been previously tagged with the desired result.In supervised learning, the training data provided to the computers act as the administrator, instructing them on how to accurately anticipate the outcome. It used the same philosophy that a student would use when learning under a teacher's guidance.

The process of providing the machine learning model with accurate input and output data is known as supervised learning. Finding a mapping function to connect the input variable (x) with the output variable (y) is the goal of supervised learning algorithms. Supervised learning may be used in the real world for a variety of tasks, including risk assessment, picture categorization, fraud detection, spam filtering, etc.

XGBOOST

The grade boosted trees technique is widely used and successfully implemented in open-source software by XGBoost. Grade boosting is a supervised learning method that combines the estimate of a number of weaker, simpler models in an effort to predict a target variable directly.Large dataset performance, usability, and speed are all priorities in the design of XGBoost.It doesn't need parameter optimisation or tweaking, thus it may be used immediately after installation without requiring any more settings.

CNN ALGORITHMS

Convolutional neural networks (CNNs) are a class of deep learning algorithms that excel in processing and recognising images. Convolutional layers, pooling layers, and totally linked layers are among the many layers that make up this structure.

Convolutional layers, which apply filters to the input picture to highlight characteristics like edges, textures, and forms, are an important component of a CNN. The output of the convolutional layers is further sent via pooling layers, which are used to down-test the point maps and narrow the spatial scope while preserving the most crucial data. One or more totally linked layers that are used to forecast and categorise the picture are also fed via the output of the pooling layers.

MODELING

In order to get useful results, the repurposed data is employed in forecast models. Machine leaning is a technique used in this forecast modelling. Select a model that works with your data and the issue you're attempting to solve. Direct retrogression, decision trees, and neural networks are a few often used models for deal prediction. To obtain the optimal performance, train the model on the training set and optimise its hyperparameters. This might include regularising, conforming the literacy rate, or using other hyperparameters. Use colourful metrics to estimate the model's performance, such as delicacy, perfection, recall, and F1-score. This will enable you to assess the model's effectiveness and evaluate if any more improvements are possible.

To ensure that the model continues to provide correct prognostications, examiners and modernizers regularly monitor the model's performance and update it as needed. This can include updating the model parameters or gathering additional data.

RESULTS

SUMMARY OF PRODUCT CATALOGUE

Total number of Unique Products	16217
Total number of Product Categories	9
Total number of Sub-Categories	77
Total number of Product types	358
Total number of brands	1933

Model	MAE	RMSE	Score

XG Boost	0.0036	0.0608	0.9998

# VISUALISE THE SIMILAR PRODUCT RECOMMENTAION: LOAD THE VGG PRE-TRAINED MODEL FROM KERAS.

```
Model: "model_2"
_____
Layer (type)                 Output Shape              Param #
=================================================================
input_3 (InputLayer)         [(None, 224, 224, 3)]     0

block1_conv1 (Conv2D)        (None, 224, 224, 64)      1792

block1_conv2 (Conv2D)        (None, 224, 224, 64)      36928

block1_pool (MaxPooling2D)   (None, 112, 112, 64)      0

block2_conv1 (Conv2D)        (None, 112, 112, 128)     73856

block2_conv2 (Conv2D)        (None, 112, 112, 128)     147584

block2_pool (MaxPooling2D)   (None, 56, 56, 128)       0

block3_conv1 (Conv2D)        (None, 56, 56, 256)       295168

block3_conv2 (Conv2D)        (None, 56, 56, 256)       590080

block3_conv3 (Conv2D)        (None, 56, 56, 256)       590080

block3_pool (MaxPooling2D)   (None, 28, 28, 256)       0

block4_conv1 (Conv2D)        (None, 28, 28, 512)       1180160

block4_conv2 (Conv2D)        (None, 28, 28, 512)       2359808

block4_conv3 (Conv2D)        (None, 28, 28, 512)       2359808

block4_pool (MaxPooling2D)   (None, 14, 14, 512)       0

block5_conv1 (Conv2D)        (None, 14, 14, 512)       2359808

block5_conv2 (Conv2D)        (None, 14, 14, 512)       2359808

block5_conv3 (Conv2D)        (None, 14, 14, 512)       2359808

block5_pool (MaxPooling2D)   (None, 7, 7, 512)         0

flatten (Flatten)            (None, 25088)             0

fc1 (Dense)                  (None, 4096)              102764544

fc2 (Dense)                  (None, 4096)              16781312

=================================================================
Total params: 134,260,544
Trainable params: 134,260,544
Non-trainable params: 0
_____
```

FEED AN ONE IMAGE IN TO THE CNN

We first see the results that the CNN produces after processing a single picture.
The following steps are:

1. loading the image
2. preparing the image to feed it into the CNN
3. get the CNN output which will correspond to the image features.

```
image loaded successfully!
image batch size (1, 224, 224, 3)
```
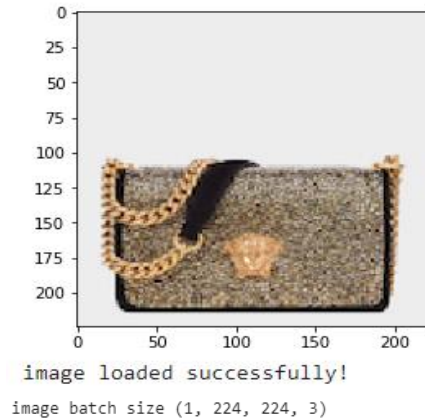
GET THE EXTRACTED FEATURES

```
1/1 [==============================] - 1s 1s/step
features successfully extracted!
number of image features: 4096
array([[0.      , 0.3754241, 0.      , ..., 0.      , 0.      ,
        0.      ]], dtype=float32)
```

FEED ALL IMAGE IN TO THE CNN

We were able to do the feature extraction process for one image. Now let's do it for all our images

GET THE EXTRACTED IMAGE

```
69/69 [==============================] - 1280s 18s/step
features successfully extracted!
(2194, 4096)
```

COMPUTE COSINE SIMILARITIES

We can calculate similarity metrics for every picture pair now that we have features for every image.We will use here the cosine similarity metric.

|	/content/drive/MyDrive/products/pro/style/5_2_060.png	/content/drive/MyDrive/products/pro/style/3_9_005.png	/c
/content/drive/MyDrive/products/pro/style/5_2_060.png	1.000000	0.333375	
/content/drive/MyDrive/products/pro/style/3_9_005.png	0.333375	1.000000	
/content/drive/MyDrive/products/pro/style/3_8_003.png	0.305859	0.488672	
/content/drive/MyDrive/products/pro/style/4_2_025.png	0.401138	0.238162	
/content/drive/MyDrive/products/pro/style/4_6_027.png	0.360251	0.344255	

5 rows × 2194 columns

RETRIEVE MOST SIMILARITIES

The last stage is to construct a function that returns the items that are aesthetically the most similar to a specified product.
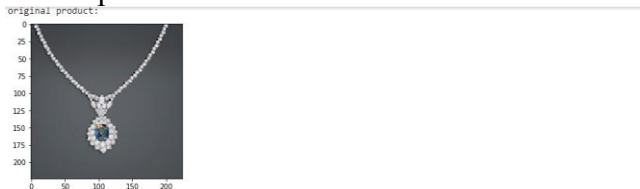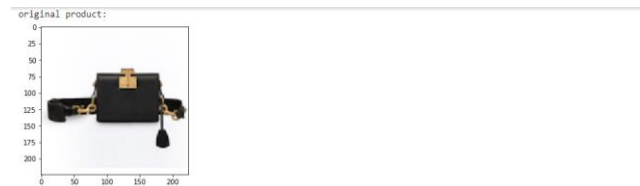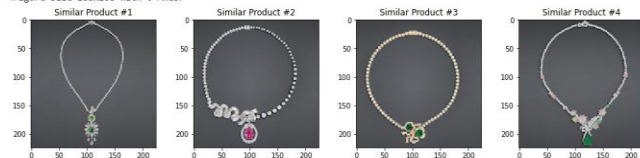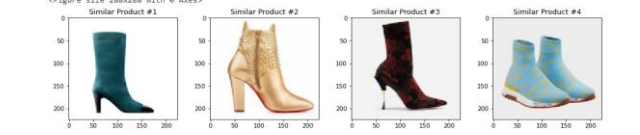
CONCLUSION

In conclusion, machine learning is having a substantial influence in two key areas: digital marketing prediction and comparable product suggestion. Machine learning algorithms may be used by organisations to create predictions and suggestions that improve their marketing tactics and boost profits by using vast amounts of customer data.

This enhances the customer experience, raises customer happiness, and helps firms close sales. In general, machine learning is a crucial tool for companies trying to enhance their digital marketing tactics and provide visitors with supported statements. Businesses may form data-driven decisions that boost profits and ensure long-term success by employing these tools to analyse customer data.

In order to anticipate sales volume and other key performance indicators (KPIs), machine learning algorithms analyse data like website business, client behaviour, and purchase history to identify important aspects. This helps companies to maximise their return on investment and optimise their marketing strategy.Similar product suggestions are made by machine learning algorithms that analyse client data to spot trends and provide tailored product recommendations.

REFERENCE

1. Thomas Ragg, Wolfram Menzel, Walter Baum, Michael Wigbers, Bayesian learning for sales rate prediction for thousands of retailers, Neurocomputing, Volume 43, Issues 1–4,2002,

2. H. Yuan, W. Xu and M. Wang, "Can online user behavior improve the performance of sales prediction in E-commerce?," 2014 IEEE International Conference on Systems, Man, and Cybernetics (SMC), San Diego, CA, USA, 2014, pp. 2347-2352.

3. Malar, P. J. M. A. J. (2016). Innovative digital marketing trends 2016. 2016 International Conference on Electrical, Electronics, and Optimization Techniques (ICEEOT).

4. Noaman M. Ali, Abdullah Alshahrani, Ahmed M. Alghamdi and Boris Novikov .Online Products Recommendations System Based on Analyzing Customers Reviews. 5. T, G., Choudhary, R., & Prasad, S. (2018). Prediction of Sales Value in online shopping using Linear Regression. 2018 4th International Conference on Computing Communication and Automation (ICCCA)

6. Punam, K., Pamula, R., & Jain, P. K. (2018). A Two-Level Statistical Model for Big Mart Sales Prediction. 2018 International Conference on Computing, Power and Communication Technologies (GUCON).

7. L. Huang, Z. Dou, Y. Hu and R. Huang, "Online Sales Prediction: An Analysis With Dependency SCOR-Topic Sentiment Model," in IEEE Access, vol. 7, pp. 79791-79797, 2019,

8. Chen, T., Yin, H., Chen, H. et al. Online sales prediction via trend alignment-based multitask recurrent neural networks. Knowl Inf Syst 62, 2139–2167 (2020).

9. J Sekban - 2019.Applying machine learning algorithms in sales prediction.

10. Michael Giering.Retail sales prediction and item recommendations using customer demographics at store level

11. Sharma, A. K., Goel, N., Rajput, J., & Bilal, M. (2020). An Intelligent Model for Predicting the Sales of a Product. 2020 10th International Conference on Cloud Computing, Data Science & Engineering (Confluence)

12. Mastan vali, K.Sankeerthana, B. Naveen, N.Vishal. PREDICTION OF ONLINE SALES USING LINEAR REGRESSION (IJCRT), Hydrabad, India.

13. S. K. Punjabi, V. Shetty, S. Pranav and A. Yadav, "Sales Prediction using Online Sentiment with Regression Model," 2020 4th International Conference on Intelligent Computing and Control Systems (ICICCS), Madurai, India, 2020, pp. 209-212

14. P, R., & M, S. (2021). Predictive Analysis for Big Mart Sales Using Machine Learning Algorithms. 2021 5th International Conference on Intelligent Computing and Control Systems (ICICCS).

15. Niu, Y. (2020). Walmart Sales Forecasting using XGBoost algorithm and Feature engineering. 2020 International Conference on Big Data & Artificial Intelligence & Software Engineering (ICBASE).

16. Singh, B., Kumar, P., Sharma, N., & Sharma, K. P. (2020). Sales Forecast for Amazon Sales with Time Series Modeling. 2020 First International Conference on Power, Control and Computing Technologies (ICPC2T).

17. Purvika Bajaj, Renesa Ray, Shivani Shedge, Shravani Vidhate, Prof. Dr. Nikhilkumar Shardoor. SALES PREDICTION USING MACHINE LEARNING ALGORITHMS (IRJET).

18. P, R., & M, S. (2021). Predictive Analysis for Big Mart Sales Using Machine Learning Algorithms. 2021 5th International Conference on Intelligent Computing and Control Systems (ICICCS)