IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

"Tiles And Marbles E-Commerce Inventory Management Through Sales Prediction"

Dr. Priya Nandihal. Associate Professor, Department of Computer Science and Engineering.
Palanati Krishna Hanuman, Raman Singh, Rishav Kumar, Rohan Ashtaputre.
UG Students, Department of Computer Science and Engineering.
Dayananda Sagar Academy of Technology and Management, Bangalore, Karnataka India

Abstract

This paper explores the intersection of predictive analytics and inventory management in the tiles and marbles e-commerce sector. Focused on developing an innovative sales prediction model, the study aims to revolutionize demand forecasting, optimize inventory levels, enhance marketing campaigns, and ultimately reduce costs. The investigation addresses the limitations of existing solutions, seeking to empower businesses with a tailored inventory management system to gain a competitive edge in this dynamic industry.

1. Introduction:

In the dynamic world of e-commerce, the ability to accurately forecast product demand is paramount. This capability serves as the cornerstone for optimizing inventory management and crafting effective marketing strategies. However, the realm of product demand forecasting is a complex puzzle, with numerous variables influencing consumer behavior. To address this challenge, the development of sales prediction models emerges as a powerful solution. These models harness the wealth of historical sales data and incorporate external factors like marketing campaigns, competitive activities, and economic trends. Their purpose is clear: to predict future sales with precision and provide e-commerce platforms with the insights needed to make data-driven decisions. These models endeavors to create a robust sales prediction model, a cornerstone in enhancing inventory management, reducing costs, and elevating revenue in the competitive and ever-evolving landscape of e-commerce. The paper evaluates seven models for e-commerce sales prediction, including two linear models, three machine learning models, and two deep learning models. The models are evaluated on the basis of their performance in predicting sales for 3,049 products over 1,941 days across three U.S. states. The models are trained on the M5 data set provided by Walmart, which includes sales history, calendar information, and price information for each product. The M5 data set provided by Walmart is used to train and evaluate the sales prediction models. The data set includes sales history, calendar information, and price information for 3,049 products over 1,941 days across three U.S. states. The data set is used to evaluate the performance of seven models for e-commerce sales prediction. The paper finds that adding calendar and price information improves model performance, but complex machine learning and deep learning models do not have a significant advantage over simple linear models. The best performing model is a linear model that includes calendar and price information. The paper also finds that the performance of machine learning and deep learning models is not significantly improved by increasing the complexity of the model.[1] The paper suggests exploring the performance of machine learning and deep learning models on larger data sets and considering more indicators that may affect e-commerce sales prediction. XGBoost is another gradient boosting framework that uses decision trees and is known for its speed and accuracy. The hybrid model combines the strengths of both frameworks to improve the accuracy of sales predictions.[2] The paper emphasizes the importance of preprocessing data and feature selection for model accuracy. Feature engineering is the process of selecting and transforming raw data into features that can be used by machine learning models. The paper describes

how they used feature engineering to create new features such as lagged variables, moving averages, and rolling standard deviations. These features were then used to train the hybrid model.[3]

The integrated model outperforms individual LightGBM and XGBoost models with an RMSE of 2.07. RMSE stands for Root Mean Squared Error, which is a measure of the difference between predicted and actual values. The lower the RMSE, the better the model's performance. The paper also provides a detailed analysis of the model's performance and compares it to other models in the literature.

Accurate sales forecasting is crucial for optimizing business plans and reducing risks. The paper highlights the importance of sales forecasting and how it can be used to make informed decisions about inventory management, marketing strategies, and resource allocation.

The paper discusses the use of various machine learning algorithms for predictive analysis in retail sales, specifically for Big Mart. The techniques include Linear Regression, Polynomial Regression, Ridge Regression, and XGBoost Regression.

2. Literature Survey:

The paper discusses an efficient and accurate sales forecasting model using machine learning, specifically XGBoost[1]. The authors propose a novel approach to sales forecasting that uses XGBoost, a popular machine learning algorithm, to predict future sales based on historical sales data. The model is designed to be both efficient and accurate, making it ideal for use in retail settings where sales forecasting is critical[1].

The paper highlights the importance of feature engineering in extracting features from historical sales data for better forecasting accuracy[2]. The authors explain how they used a variety of feature engineering techniques to extract relevant features from the Walmart retail goods dataset, including time-based features, holiday-based features, and more. By carefully selecting and engineering these features, the authors were able to improve the accuracy of their sales forecasting model.

The paper emphasizes XGBoost's scalability and efficiency, running ten times faster than other machine learning algorithms with fewer resources[3]. The authors explain how XGBoost is able to achieve such high levels of scalability and efficiency, highlighting its use of parallel processing and distributed computing. They also note that XGBoost is highly customizable, allowing users to fine-tune the algorithm to their specific needs.

The proposed model was tested on a Walmart retail goods dataset from Kaggle, demonstrating superior performance in sales prediction with a lower RMSSE score compared to other methods. The authors provide detailed experimental results showing the effectiveness of their proposed model, including a comparison with other state-of-the-art sales forecasting methods[4]. The results demonstrate that the proposed model outperforms other methods in terms of both accuracy and efficiency, making it a promising approach for sales forecasting in retail settings.[4]

The paper presents a study for the prediction of e-commerce business market growth using machine learning algorithm[1]. The paper aims to help online vendors manage their inventories and customer retention by analyzing the sales data of an e-commerce company and forecasting the income and demand of different products per quarter[5]. The paper also provides analysis results and insights on the most sold commodities and their frequencies of purchase per quarter. Online venders of business market manage their inventories on virtual prediction bases for full filling the basic need of demand-supply chain of customers. Authorizing traditional ways and analysis methods are not ensuring the rate of reliability of the sales prediction. To produce more precise predictions and analysis, we use ML algorithm. In this paper, we utilized the selling data set of an E-commerce company and segregated it, in different quarters then calculating the sale income per quarter. After that we divided the dataset in the proportion of 70% and 30% for Training data set and Testing data set. By applying machine learning algorithm, we will be predicting income of next quarters as well as analysis the maximally sold commodities with their frequencies of purchase per quarter. Then provide analysis results and prediction of customer’s purchase patterns to the business organization to make a strategy to take a competitive advantage by sustaining and accumulating for their goods management and planning for inventories. [5]

The paper examines how manufacturing firms’ inventory investment decisions vary depending on the direction of sales changes[1]. It proposes that managers respond asymmetrically to sales increases versus decreases due to factors such as production smoothing, expectations of future demand, and stockout costs[2]. It also explores the implications of asymmetric inventory investment for forecasting future sales growth and earnings[34].[6]

[6]We validate this claim by showing that managers' expectations of future demand and desire to avoid inventory stockouts are important determinants of this asymmetry. In addition, we find that asymmetric inventory investment provides useful information for predicting future sales growth, and that both managers' and analysts' sales forecasts are positively associated with the asymmetry[6]

[7]The paper studies the relationship between sales surprise and inventory turnover using two types of sales forecasts: Holt's method and management forecasts. Sales surprise is the ratio of actual sales to forecasted sales, and inventory turnover is the ratio of the cost of goods sold to average inventory[1]. The paper uses data from Japanese listed companies in manufacturing and retail industries for the period 1997-2014.[7]

The paper proposes a model to predict the sales quantity of multiple products by adopting the Recency-Frequency-Monetary (RFM) concept and Fuzzy Analytic Hierarchy Process (FAHP) method1.[8] The paper aims to improve the inventory management and customer satisfaction of a company selling online.[8] The paper uses a case study of a digital content provider to illustrate the application of the proposed model.[8] The paper evaluates the prediction accuracy of the model using the mean absolute percentage error (MAPE) and compares it with other methods2.[8]

Predicting excessive or inaccurate sales can lead to inventory-related cost issues, resulting in inefficient investment. Therefore, it takes an economic model that can be used to predict the number of sales in order that inventory costs become more efficient, service to consumers could be improved and provide a significant competitive advantage for the company. To answer these problems, this study will propose to construct a model that can predict the sales quantity by adopting the RFM concept. Recency, frequency and monetary (RFM) were a powerful and well-known concept in database marketing, and were widely used to measure the value of customers based on their prior purchasing history. The RFM concept had also been successfully integrated into the mining process in the past few years [1-3], giving rise to the idea that the RFM concept also can be applied to predict sales based on prior sales history.[8]

The paper proposes a deep learning approach for predicting the sales of retail stores using point-of-sale (POS) data12. The paper uses three years of POS data from supermarkets in Japan and constructs a binary classification model that predicts whether the sales of the next day will increase or decrease compared to the current day. The paper compares the performance of the deep learning model with a logistic regression model and evaluates the models using various metrics such as accuracy, precision, recall, and F-measure.[9]

The paper presents a novel approach to explain the predictions of black-box machine learning models in the context of business-to-business (B2B) sales forecasting1. The authors use two general explanation methods, EXPLAIN and IME, to generate uniform and interactive explanations for any ML model. They demonstrate the applicability and usefulness of their approach on a real-world case of complex B2B sales forecasting, where they use various ML techniques to build prediction models and compare them. They also provide a publicly available data set to encourage further research in this domain2.[10]

3. **Limitations**

After going through the papers in details we found few of the drawbacks. These are:

Paper TITLE & authors (IEEE format citation)	Concept	Limitation	Tool used/ specific algorithm / platform (Just name)
[1] "Zixuan Huo, "Sales Prediction based on Machine Learning" 2nd International Conference on E-Commerce and Internet Technology (ECIT) 2021."	This paper evaluates two linear models, three machine learning models, and two deep learning models.	The paper's limitations include a small dataset, lack of certain input features, limited model selection, and potentially inappropriate evaluation metrics. It uses a dataset of 5.4 years of sales history for 3049 products in 10 stores, which may not capture long-term trends and seasonal patterns[1]. The paper only considers the first type of product or merchant attribute characteristics and does not include the second type of product reviews and derivative indicators, or the third type of product online search information. It compares two linear models, three machine learning models, and two deep learning models, but there may be other more suitable models[23]. The paper uses the model training time and Root Mean Squared Error (RMSE) as the evaluation metrics, which may not reflect the practical value of the models for e-commerce businesses[4]. Other metrics may be more appropriate.	Machine Learning, Deep Learning
[2] "Jingru Wang, "A hybrid machine learning model for sales prediction", International Conference on Intelligent Computing and Human-Computer Interaction (ICHCI) 2020."	This paper is based on the LightGBM framework and the XGBoost framework to build a sales forecast model.	The paper's limitations include a small dataset, lack of certain input features, limited model selection, and potentially inappropriate evaluation metrics. It uses a dataset of 5.4 years of sales history for 3049 products in 10 stores, which may not capture long-term trends and seasonal patterns. The paper only considers the first type of product or merchant attribute characteristics and does not include the second type of	LightGBM, XGBoost

		product reviews and derivative indicators, or the third type of product online search information. It compares two linear models, three machine learning models, and two deep learning models, but there may be other more suitable models. The paper uses the model training time and Root Mean Squared Error (RMSE) as the evaluation metrics, which may not reflect the practical value of the models for e-commerce businesses. Other metrics may be more appropriate.	
[3] "Ranjitha P, Spandana M, "Predictive Analysis for Big Mart Sales Using Machine Learning Algorithms", Fifth International Conference on Intelligent Computing and Control Systems (ICICCS), 2021."	A predictive model was developed using Xgboost, Linear regression, Polynomial regression, and Ridge regression techniques for forecasting the sales of a business.	The paper's limitations include a small dataset, lack of certain input features, limited model selection, and potentially inappropriate evaluation metrics. It uses a dataset of 5.4 years of sales history for 3049 products in 10 stores, which may not capture long-term trends and seasonal patterns. The paper only considers the first type of product or merchant attribute characteristics and does not include the second type of product reviews and derivative indicators, or the third type of product online search information. It compares two linear models, three machine learning models, and two deep learning models, but there may be other more suitable models. The paper uses the model training time and Root Mean Squared Error (RMSE) as the evaluation metrics, which may not reflect the practical value of the models for e-commerce businesses. Other metrics may be more appropriate.	Polynomial regression, Ridge regression
[4] "Xie dairu, and Zhang Shilong, "Machine Learning Model for Sales Forecasting by Using	This paper proposes an efficient and accurate sales forecasting model using machine learning.	The paper's limitations include a small dataset, lack of certain input features, limited model selection, and potentially inappropriate evaluation metrics. It uses a dataset of 5.4	XGBoost

XGBoost" IEEE International Conference on Consumer Electronics and Computer Engineering (ICCECE) 2021."		years of sales history for 3049 products in 10 stores, which may not capture long-term trends and seasonal patterns. The paper only considers the first type of product or merchant attribute characteristics and does not include the second type of product reviews and derivative indicators, or the third type of product online search information. It compares two linear models, three machine learning models, and two deep learning models, but there may be other more suitable models. The paper uses the model training time and Root Mean Squared Error (RMSE) as the evaluation metrics, which may not reflect the practical value of the models for e-commerce businesses. Other metrics may be more appropriate.	
[5] "Shilpi Kulshrestha and M. L. Saini, "Study for the Prediction of E-Commerce Business Market Growth Using Machine Learning Algorithm" 5 th IEEE International Conference on Recent Advances and Innovations in Engineering- ICRAIE 2020."	Predicting the e-commerce business market growth using machine learning algorithms1.	The paper uses a single dataset of an e-commerce website, which may not be representative of the general online market and customer behavior. The paper applies only one machine learning algorithm, Holt's Linear Trend Model, which may not be the best or the most accurate method for forecasting sales values. The paper does not consider other factors that may affect the sales performance, such as competitors, customer feedback, marketing strategies, product quality, etc. The paper does not evaluate the reliability or validity of the predicted values, nor does it compare them with other existing methods or models.	Holt's Linear Trend Model, a machine learning algorithm based on time series analysis
[6] "Iny Hwang, Taejin Jung,	Asymmetric inventory management,	The document is not peer-reviewed. The document is an	Asymmetric inventory investment decisions

Woo-Jong Lee and Daniel G. Yang, "Asymmetric Inventory Management and the Direction of Sales Changes" July 2020."	Determinants of asymmetry, Implications of sales forcast	accepted article that has not been through the copyediting, typesetting, pagination and proofreading process[1]. This means that it may contain errors or inaccuracies that could affect its validity and reliability. The document uses a linear inventory model. The document assumes a linear relation between inventory investment and sales changes, which may not capture the complex and dynamic nature of inventory management[2]. A nonlinear or stochastic inventory model may be more appropriate for analyzing inventory behavior in response to demand uncertainty and stockout costs. The document focuses on manufacturing firms. The document limits its sample to manufacturing firms with SIC codes between 2000 and 3999, which may limit the generalizability of its findings to other industries or sectors[3]. The document may need to consider the differences in inventory management practices and implications across different types of firms, such as service firms, retailers, or wholesalers.	
[7] "Hiroki Sano and Kazuo Yamada, "Prediction accuracy of sales surprise for inventory turnover" International Journal of Production Research 2020."	The document is about how sales surprise affects inventory turnover	The paper only focuses on manufacturing and retail industries in Japan, which may limit the generalisability of the findings to other sectors or countries. The paper uses point estimates of management forecasts, which may not reflect the uncertainty or range of possible outcomes that managers consider when making their forecasts. The paper relies on publicly available data, which may not capture the true sales forecasts	Holt's method and management forecasts.

		or inventory management practices of firms that are not disclosed or reported accurately. The paper does not account for other factors that may affect inventory turnover, such as product life cycle, product variety, demand uncertainty, or supply chain coordination.	
[8] "Rendra Gustriansyah, Dana Indra Sensuse, Arief Ramadhan, "A Sales Prediction Model Adopted the Recency Frequency-Monetary Conce" Indonesian Journal of Electrical Engineering and Computer Science 2017."	The document is a research paper that proposes a model to predict the sales quantity of pharmaceutical products in a pharmacy using the RFM (Recency-Frequency-Monetary) concept and the FAHP (Fuzzy Analytic Hierarchy Process) method.	**Outdated information**: The web page context is a preprint version of a paper that was accepted for publication in 2016. Some of the information and references may be outdated or superseded by newer research. **Incomplete manuscript**: The web page context only contains part of the original paper, namely the introduction and the point-to-point responses to reviewers. The main content of the paper, such as the methods, results, discussion, and conclusion, are not included in the context. **Specific domain**: The web page context is focused on a specific domain of explaining machine learning models in sales predictions[1]. The terminology and concepts may not be familiar or relevant to a general audience or other domains. **Limited scope**: The web page context only presents one approach to building an intelligent system with general explanation methodology. It does not compare or contrast its approach with other existing or alternative methods. It also does not provide empirical evidence or evaluation of its effectiveness and usability.	RFM concept, FAHP method,

[9] "Yuta Kaneko, Katsutoshi Yada, "A Deep Learning Approach for the Prediction of Retail Store Sales" 16th International Conference on Data Mining Workshops 2016."	The study applies deep learning to construct a sales prediction model using POS data from a retail business, demonstrating the potential of deep learning in practical applications.	**Deep learning for retail sales prediction**: The web page is a research paper that proposes a method to forecast the sales of a supermarket using deep learning, a machine learning technique that can handle large-scale and complex data. **Data set and model construction**: The paper uses three years of point-of-sale (POS) data from supermarkets in Japan, and constructs a binary classification model that predicts whether the sales of the next day will increase or decrease based on the sales of the current day[1]. The paper also compares different categories of product attributes, ranging from 62 to 3312, and applies regularization techniques to improve the model performance. **Experiments and results**: The paper evaluates the deep learning model using various indicators, such as accuracy, precision, recall, and F-measure, and compares it with a logistic regression model. The paper shows that the deep learning model achieves higher accuracy than the logistic regression model, especially when the number of product attributes is large. The paper also shows that the deep learning model is robust to the sparsity of the data, and that L1 regularization is effective for enhancing the predictive accuracy[2]. **Conclusion and future work**: The paper concludes that deep learning is a useful technique for sales prediction in retail businesses, and that it can provide valuable information for store managers to plan marketing strategies[3]. The paper also suggests that the accuracy of the model can be further improved by optimizing	The study uses deep learning models with L1 regularization for constructing the sales prediction model.

		the number of attributes and introducing additional regularization[4]. The paper also states that the next step is to predict the sales figures themselves, rather than just the increase or decrease in sales[5].	
[10] "Marko Bohanec, Mirjana Kljajic Bor´stnar, Marko Robnik Sikonja, "Explaining machine learning models in sales predictions, Expert Systems With Applications" 2016."	The project is about creating an intelligent system that can explain machine learning models for sales forecasting	**Outdated information**: The web page context is a preprint version of a paper that was accepted for publication in 2016. Some of the information and references may be outdated or superseded by newer research. **Incomplete manuscript**: The web page context only contains part of the original paper, namely the introduction and the point-to-point responses to reviewers. The main content of the paper, such as the methods, results, discussion, and conclusion, are not included in the context. **Specific domain**: The web page context is focused on a specific domain of explaining machine learning models in sales predictions[1]. The terminology and concepts may not be familiar or relevant to a general audience or other domains. **Limited scope**: The web page context only presents one approach to building an intelligent system with general explanation methodology. It does not compare or contrast its approach with other existing or alternative methods. It also does not provide empirical evidence or evaluation of its effectiveness and usability.	

Conclusion:

After going through the reference papers for the project things can be concluded that there isn't a proper inventory management tool for the Tiles and Marbels sector. The technologies used could be refined to a better extent with better possibilities. The market's versatility is too high hence we need a large data of real-world value to train the machine for proper results. Improvement needs to be done in a fast pace to overcome the rate at which changes are happening in the real-world scenarios.

Therefore, there will be a need for a dedicated inventory management tool in tiles and marbles sector. A tool that can read the past sales data and can predict future sales probabilities for the growth and sustainment of the shop in the market of changing trends.

References:

[1] "Zixuan Huo, "Sales Prediction based on Machine Learning" 2nd International Conference on E-Commerce and Internet Technology (ECIT) 2021." Doi: 0.1109/ECIT52743.2021.00093

[2] "Jingru Wang, "A hybrid machine learning model for sales prediction", International Conference on Intelligent Computing and Human-Computer Interaction (ICHCI) 2020." Doi: 10.1109/ICHCI51889.2020.00083

[3] "Ranjitha P, Spandana M, "Predictive Analysis for Big Mart Sales Using Machine Learning Algorithms", Fifth International Conference on Intelligent Computing and Control Systems (ICICCS), 2021." Doi: 0.1109/ICICCS51141.2021.9432109

[4] "Xie dairu, and Zhang Shilong, "Machine Learning Model for Sales Forecasting by Using XGBoost" IEEE International Conference on Consumer Electronics and Computer Engineering (ICCECE) 2021." Doi: 0.1109/ICCECE51280.2021.9342304

[5] "Shilpi Kulshrestha and M. L. Saini, "Study for the Prediction of E-Commerce Business Market Growth Using Machine Learning Algorithm" 5 th IEEE International Conference on Recent Advances and Innovations in Engineering- ICRAIE 2020." Doi: 0.1109/ICRAIE51050.2020.9358275

[6] "Iny Hwang, Taejin Jung, Woo-Jong Lee and Daniel G. Yang, "Asymmetric Inventory Management and the Direction of Sales Changes" July 2020." Doi: 10.1111/1911-3846.12635

[7] "Hiroki Sano and Kazuo Yamada, "Prediction accuracy of sales surprise for inventory turnover" International Journal of Production Research 2020." Doi: 10.1080/00207543.2020.1778205

[8] "Rendra Gustriansyah, Dana Indra Sensuse, Arief Ramadhan, "A Sales Prediction Model Adopted the Recency Frequency-Monetary Conce" Indonesian Journal of Electrical Engineering and Computer Science 2017." Doi: 10.11591/ijeecs.v6.i3.pp711-720

[9] "Yuta Kaneko, Katsutoshi Yada, "A Deep Learning Approach for the Prediction of Retail Store Sales" 16th International Conference on Data Mining Workshops 2016." Doi: I 10.1109/ICDMW.2016.154

[10] "Marko Bohanec, Mirjana Kljajic Bor´stnar, Marko Robnik Sikonja, "Explaining machine learning models in sales predictions, Expert Systems With Applications" 2016." Doi: 10.1016/j.eswa.2016.11.010