International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 11s
DOI: https://doi.org/10.17762/ijritcc.v11i11s.8170
***Article Received:** 24 June 2023 **Revised:** 24 August 2023 **Accepted:** 09 September 2023*

Customer Segmentation and Business Sales Forecasting using Machine Learning for Business Development

Pravin Malviya[1], Dr. Vijay Bhandari[2] Pankaj Singh Sisodiya[3], Saurabh Suman[4]
[1,3,4]Research Scholar Department of Computer Science & Engineering , MPU Bhopal, India.
[2]Associate Professor Department of Computer Science & Engineering , Bhopal, India.

Abstract: This study explores the application of machine learning techniques for business development, focusing on sales prediction and customer segmentation, using a Walmart dataset. Performance metrics include Mean Absolute Error (MAE) and R2 scores. Our hybrid approach combines the BIRCH algorithm with time-lagged machine learning (TL-ML). The results reveal that customer segmentation significantly improves model performance across all metrics. Among the techniques tested, models incorporating customer segmentation (CS-RFR and CS-TL-ML) outperform standard Random Forest Regressor models. Specifically, CS-TL-ML shows a slight advantage in terms of both lower MAE and higher R2 scores, confirming its efficacy for sales prediction and customer segmentation tasks.

Keywords: Customer Segmentation, Sales Prediction, BIRCH Algorithm, Time-Lagged Machine Learning, Business Development.

I. Introduction

In the era of digital transformation, Machine Learning (ML) has emerged as a disruptive technology that is profoundly altering the landscape of business development. As companies across industries strive to stay ahead of the curve, the integration of machine learning technologies offers unprecedented advantages. From making business processes more streamlined to enhancing decision-making through data-driven insights, machine learning is rapidly becoming an essential tool in the contemporary business toolkit. Beyond the broad applications in fields like healthcare, finance, and autonomous vehicles, ML's influence has become increasingly pervasive in business development, enabling companies to adopt more strategic and data-centric approaches to growth [1]. Businesses today amass enormous volumes of customer data, from purchasing habits to engagement metrics. Traditional methods of data analysis often fall short in the task of interpreting this wealth of information in meaningful ways. Machine learning algorithms, however, can rapidly and efficiently segment customers based on an array of metrics such as behaviour, purchasing patterns, and other key performance indicators. This segmentation not only allows for more effective targeted marketing but also enhances customer engagement by delivering more personalized experiences [2][3]. In the fiercely competitive market, the ability to predict future sales trends is indispensable for sustainable growth. Machine learning leverages historical data to make accurate predictions about which leads are most likely to convert into customers. This predictive analytics capability empowers businesses to allocate their resources more efficiently and focus their efforts on the most promising opportunities, thus maximizing returns on investment [4][5]. Product recommendation systems are another vital aspect where machine learning shows its prowess. By analyzing customer browsing history, purchase behaviour, and preferences, ML algorithms can provide highly tailored product suggestions that are likely to resonate with the individual consumer. This feature not only enhances the user experience but also increases the chances of successful sales conversions and customer retention [6]. By weaving these features into a cohesive machine learning-based business development module, this paper aims to provide a detailed exploration of each, while also discussing the limitations, ethical considerations, and future directions for the research. Through empirical analysis and case studies, we aspire to demonstrate how the adoption of machine learning technologies can significantly augment business development efforts, leading to more strategic and informed decisions for organizations in the 21st century [7][8].

II. Related Work

Shao et al. [1] focus on Business Intelligence (BI) in corporate finance. They propose an IoT-based Efficient Data Visualization Framework (IoT-EDVF) to manage data quality, analyze multiple data sources, and mitigate data leaks. The framework has shown promising results, such as 29.42% improvement over existing models. Gehlot et al. [2] explore the use of machine learning in telecommunications. They emphasize the role of analytics to understand mobile users' needs, particularly focusing on the effective utilization of big data analytics in cellular networks. Ghazal et al. [3] address

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 11s
DOI: https://doi.org/10.17762/ijritcc.v11i11s.8170
***Article Received:** 24 June 2023 **Revised:** 24 August 2023 **Accepted:** 09 September 2023*

Supply Chain Collaboration, advocating for machine learning to enhance decision-making processes. Their model provides strong computing ability and accurate predictions, emphasizing the importance of shared information. Huber et al. [4] employ machine learning to assess the impact of discount rates on train ticket buying behavior. They found that a one-percentage-point increase in the discount rate resulted in a 0.16 percentage point increase in rescheduled trips. Irfan et al. [5] apply artificial intelligence in food industry sales predictions. Their model aims to maximize food business profits and has shown a low mean squared error and low variance in predictions. Hawkins et al. [6] delve into metaverse live shopping analytics, emphasizing the role of data visualization tools to enhance immersive retail experiences in virtual worlds. Patriarca et al. [7] explore the use of Business Intelligence and Machine Learning in Air Navigation Service Providers. Their data-driven framework aims to elevate safety standards. Omri et al. [8] discuss project management through BI. They offer a model to facilitate team communication, planning, forecasting, and documentation. Watson et al. [9] examine the virtual economy in the metaverse. They highlight the role of customer engagement tools and predictive analytics in enhancing business competitiveness. Shewale et al. [10] focus on Home Energy Management Systems (HEMS) to optimize residential energy consumption. They survey various demand-side management strategies and suggest future research directions. Niu et. al. [11] proposed an Optimized Data Management framework using Big Data Analytics to improve organizational effectiveness, focusing on tackling challenges like plan failure and risk management. Tamang et. al. [12] discussed the benefits and challenges of combining machine learning with business intelligence, highlighting real-time data analysis and cybersecurity. Romero et. al. [13] highlighted the importance of Industry 4.0 in driving innovation and rapid responses in dynamic markets, focusing on interconnectivity and machine learning. Ranjan et. al. [14] focused on Competitive Intelligence and Big Data, presenting a comprehensive look into how organizations are currently utilizing these technologies. Patriarca et. al. [15] used Business Intelligence and Machine Learning to analyze industrial incidents, aiming to reduce their frequency and severity. Zhang et. al. [16] provided a bibliometric study of the current landscape of information systems research, with a focus on Big Data analytics and machine learning. Annapurani et. al. [17] discussed the integration of BI tools like Microsoft Power-BI in healthcare, specifically focusing on data from Electronic Health Records (EHR). Fraihat et. al. [18] focused on the Jordanian real estate market, utilizing BI and ML to provide reliable analytics and decision support for potential investors. Khan et. al. [19] proposed a demand-side management framework for smart grids, using AI algorithms for optimizing energy usage behavior and pricing. Silva et. al. [20] conducted a systematic literature review on the application of Business Analytics within Industry 4.0, identifying current practices and future research opportunities.

III. Proposed Methodology

In this section, to explore the business development following flowchart is prepared as presented in figure 1.

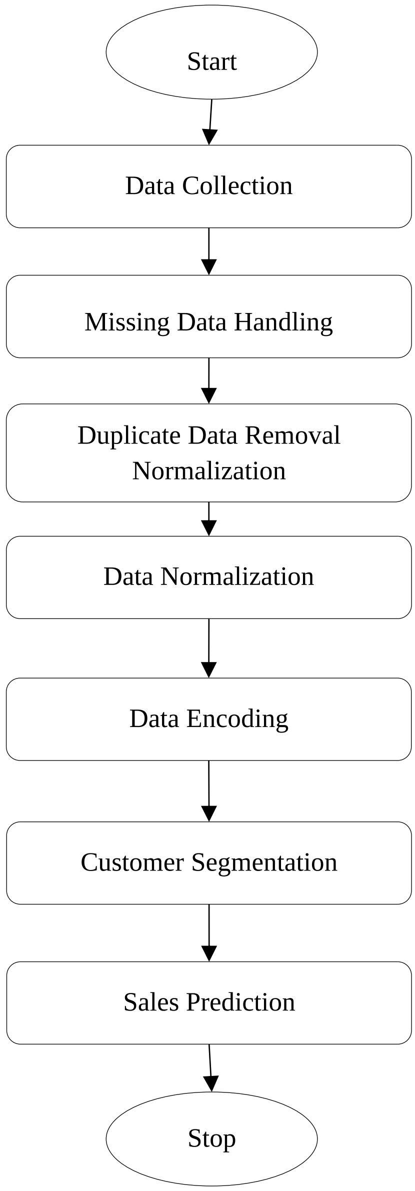


Figure 1: Proposed Flowchart

In this work, for business development we have taken a walmart dataset [21] for business development with key features of sales prediction on the basis of customer segmentation and product recommendation system. For this we have used the machine learning approach. For this we have implemented a hybrid approach of BIRCH algorithm with time-lagged machine learning for sales prediction with customer segmentation.

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 11s
DOI: https://doi.org/10.17762/ijritcc.v11i11s.8170
***Article Received:** 24 June 2023 **Revised:** 24 August 2023 **Accepted:** 09 September 2023*

Customer segmentation is the practice of dividing a company's customers into groups that reflect similarity among customers in each group. The benefit of segmenting customers is to cater more precisely to the specific needs of each segment. For this purpose, we've employed the BIRCH (Balanced Iterative Reducing and Clustering using Hierarchies) algorithm. BIRCH is particularly well-suited for very large datasets like that of Walmart. Its hybrid approach ensures efficient clustering by generating a concise representation of data. Sales prediction is integral for any business to anticipate revenue and adjust strategies accordingly. In this section, instead of building models from scratch, we've leveraged the power of with time-lagged machine learning. This method generally saves time and often provides superior results, especially if the source data/domain is closely related to the target domain.

1.1 Data Collection

Data preparation is the crucial first step in any data analysis process. It involves cleaning and transforming raw data into a format that can be analyzed. This could involve tasks such as handling missing values, removing outliers, and ensuring that the data is consistent. In this step, we have referred to existing dataset taken from Kaggle [97]. For result evaluation, the Walmart Store Sales Forecasting dataset [21] was used. comprises historical sales data from 45 Walmart stores in the US, intended for weekly sales prediction. The dataset is split into training (train.csv) and test (test.csv) segments, spanning from 2010 to 2012, with sales data withheld in the test set. Sample of dataset is presented below in figure 2.

	Store	Date	Temperature	Fuel_Price	CPI	Unemployment	IsHoliday	Type	Size
0	1	2010-02-05	42.31	2.572	211.096358	8.106	False	A	151315
1	1	2010-02-12	38.51	2.548	211.242170	8.106	True	A	151315
2	1	2010-02-19	39.93	2.514	211.289143	8.106	False	A	151315
3	1	2010-02-26	46.63	2.561	211.319643	8.106	False	A	151315
4	1	2010-03-05	46.50	2.625	211.350143	8.106	False	A	151315
5	1	2010-03-12	57.79	2.667	211.380643	8.106	False	A	151315
6	1	2010-03-19	54.58	2.720	211.215635	8.106	False	A	151315
7	1	2010-03-26	51.45	2.732	211.018042	8.106	False	A	151315
8	1	2010-04-02	62.27	2.719	210.820450	7.808	False	A	151315
9	1	2010-04-09	65.86	2.770	210.622857	7.808	False	A	151315

Figure 2: Dataset Sample

1.2 Encoding

In this step, the one-hot encoding method is used to convert the filtered data after statistical evaluation so that they can result in better prediction results. This method is adopted because it provides better prediction results of machine learning with those data that show no relationship to each other. After encoding data normalization is performed. The order of integers is treated as a significant characteristic by machine learning algorithms. In other words, a larger number will be interpreted as superior or more significant than a smaller number. While this is useful in some ordinal scenarios, other input data lacks a ranking for category values, which can cause problems with estimations and performance. That's when a single hot encoding comes to the rescue. Our training data becomes more valuable and descriptive with just one hot encoding, and it could be readily resized. We may more readily establish a probability for our values by utilizing numeric numbers. For our output values, we choose one hot encoding because it delivers more detailed predictions than single labels.

One-hot encoding is a process of converting categorical data variables so they can be provided to machine learning algorithms to improve predictions. One-hot encoding transforms nominal categorical data into a binary format in which each category is represented by a unique binary column. For a given categorical column with N unique categories:

Each category will be represented by a vector of size N.

For each category, one element of the vector will be "1" (indicating the presence of the category) and the remaining elements will be "0" (indicating the absence).

Mathematically:

Let's say we have a categorical column C that has values from the set $S = \{s_1, s_2, \ldots s_N\}$

For each value of s_i in S, the one-hot encoded vector will be: $V(si) = [0,0, \ldots ,1, \ldots ,0]$

Where,

The size of vector $V(s_i) = N$

There is a "1" at the i^{th} position and "0" in all other positions.

Algorithm 1: One Hot Encoding

Start

Input: Categorical data column.

Identify unique categories in the column and count them. Let this count be N.

For each unique category s_i.

Create a new column (binary vector) of size *N*.

Set the i^{th} position to "1" and all other positions to "0".

Replace the original categorical column with *N* new binary columns.

Output: Dataset with one-hot encoded columns.

End

This process ensures that the categorical values are represented in a format that most machine learning models can use

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 11s
DOI: https://doi.org/10.17762/ijritcc.v11i11s.8170
***Article Received:** 24 June 2023 **Revised:** 24 August 2023 **Accepted:** 09 September 2023*

effectively. However, one potential downside of one-hot encoding is the increase in the dimensionality of the dataset, which may require more memory and can also lead to the curse of dimensionality in certain algorithms.

1.3 Data Imbalance Handling

In the domain of machine learning, data imbalance is a major concern. Imbalanced data refers to datasets in which the target class has an unequal distribution of observations, i.e. one class label has a large number of observations while the other has a small number. Imbalanced data sets are a kind of classification issue in which the distribution of classes is not uniform. Traditional ways of removing data imbalance include logistic regression, up-sampling, down-sampling, and over-sampling. This approach is less accurate than PCA-based data imbalance elimination. PCA may be used to reduce a high-dimensional point to a lower-dimensional point, and then filters can be used to rank the relevance of the chosen features. The variance-covariance structure of a group of variables is described by principal component analysis (PCA) in terms of fewer new variables that are linear combinations of the original variables. The additional variables may be readily produced via Eigen analysis of the original data's covariance matrix or correlation matrix. It is advisable to do PCA on the sample correlation matrix if the variables are assessed on scales with widely differing ranges or if the units of measurement are not comparable. Finally, the PCA provides a new PC transform that is constructed by utilizing the data correlation matrix to select the best PCs among all of the features. Therefore, after encoding of data, PCA is applied to them to handle their class imbalance issue.

1.4 Customer Segmentation

For customer segmentation, BIRCH algorithm is used. The BIRCH (Balanced Iterative Reducing and Clustering using Hierarchies) algorithm is a popular method designed for clustering large datasets, particularly in situations where the dataset is too large to fit into main memory. The primary idea behind BIRCH is to build a tree structure, the CF-tree, that provides a summary of the data. The description of the BIRCH algorithm are provided:

Loading Phase: This is the first phase where BIRCH scans the dataset, ensuring that it remains within memory limits. During this phase:

An initial in-memory CF-tree (Clustering Feature tree) is constructed.

Global Clustering Phase: After having the CF-tree representation, this phase groups all the leaf entries together, even if they lie in different nodes of the CF-tree. The idea here is to perform a global or semi-global clustering. This results in a clustering that might not be highly granular, but it gives a general perspective of data distribution.

Optional Condensing Phase: Here, the CF-tree is further refined densely packed subclusters, which might have been represented separately earlier, are now merged to form larger clusters. This phase is all about making the CF-tree more compact and concise.

Optional Refining Phase: This is the last phase that primarily focuses on the precision of clusters. Data points are redistributed based on their proximity to cluster centroids. The centroids from the clusters formed in Phase 3 are used as seeds. Using these seeds, a new set of clusters is formed, ensuring that data points move closer to their most appropriate cluster.

In essence, BIRCH is designed to handle large datasets by providing a hierarchical clustering method that not only represents data efficiently but also ensures that clusters are of high quality. The multi-phase approach helps in reducing computational overhead while retaining the essence of data distribution.

Algorithm 2: Customer Segmentation (BIRCH)

Input

D= $\{t1, t2, t3 \ldots \ldots\ tn\}$ // Set of customers

T // Threshold for CF tree construction

Output

K // Set of clusters.

BIRCH clustering algorithm:

For each $ti \in D$ do

Determine correct leaf node for ti insertion;

If threshold condition is not violated then

Add ti to cluster and update CH triples;

else

if room to insert ti then

insert ti as singe cluster and update CF triples;

else

split leaf node and redistribute CF features;

end

1.5 Prediction

Time-lagged Random Forest (TLRF) is a method typically employed for predicting future values in time series data by incorporating lagged observations as features. In essence, the TLRF uses past observations to predict future ones, making it

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 11s
DOI: https://doi.org/10.17762/ijritcc.v11i11s.8170
***Article Received:** 24 June 2023 **Revised:** 24 August 2023 **Accepted:** 09 September 2023*

suitable for time series forecasting. Here's a general methodology for constructing a Time-lagged Random Forest:

Lagged Feature Creation: Given a time series $\{x_1, x_2, \ldots\ldots x_t, \ldots\}$ create lagged features. For example, to predict x_{t+1}, you might use $\{x_t, x_{t-1}, \ldots\ldots x_{t-k}, \ldots\}$ as features, where k is the number of lags.

Data Transformation: Convert the time series into a supervised learning dataset where each row represents the following:

Features: $\{x_{t-k}, x_{t-k+1}, \ldots\ldots x_t\}$

Target: x_{t+1}

While the Random Forest itself is an ensemble of decision trees and doesn't have a simple mathematical expression like linear regression, the essence of the time-lagged prediction can be illustrated.

Given a time series: $\{x_1, x_2, \ldots\ldots x_t\}$

A time-lagged representation with k lags for predicting x_{t+1} would be:

$$f(\{x_t, x_{t-1}, \ldots\ldots x_{t-k}\}) \rightarrow x_{t+1}$$

Where f represents the Random Forest prediction function. The exact nature of f is complex and is determined by the structure of the individual decision trees and their ensemble nature.

IV. Results and Discussions

In this section, implementation result performance parameters and their evaluation details are presented. The result is implemented and trained the models in the Keras framework with TensorFlow. The proposed model was trained on the python platform. To evaluate the performance of the proposed methodology following metrics are used:

Mean Absolute Error (MAE): It is used to represent the error between actual (A_i) and predicted (P_i) value.

$$MAE = \frac{\sum_{i=1}^{n} |P_i - A_i|}{n} \quad (1)$$

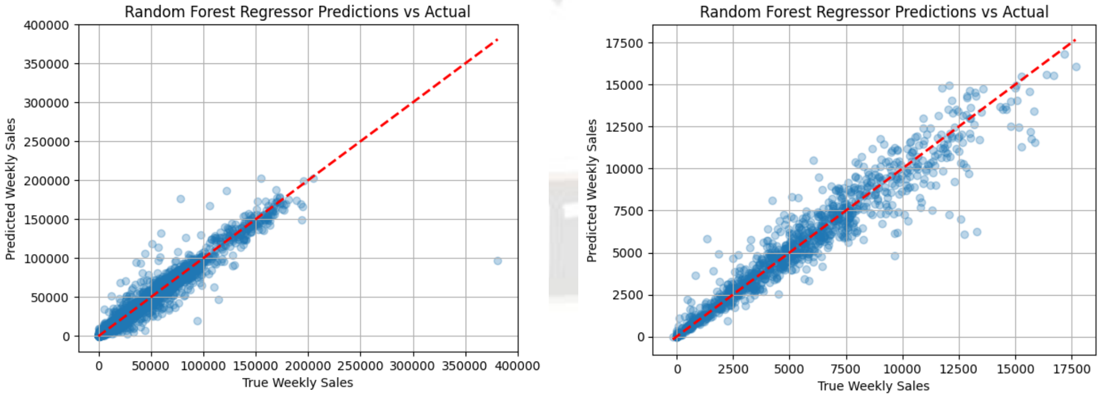
Where, n= Number of tested samples.

R-Squared: It is termed as statistical variation that is used to evaluate the variation of dependent variables with independent variables.

$$R^2 = 1 - \frac{UE_{variation}}{T_{variation}} \quad (2)$$

Where, $UE_{variation}$ = Unexplained variation and $T_{variation}$= Total variation

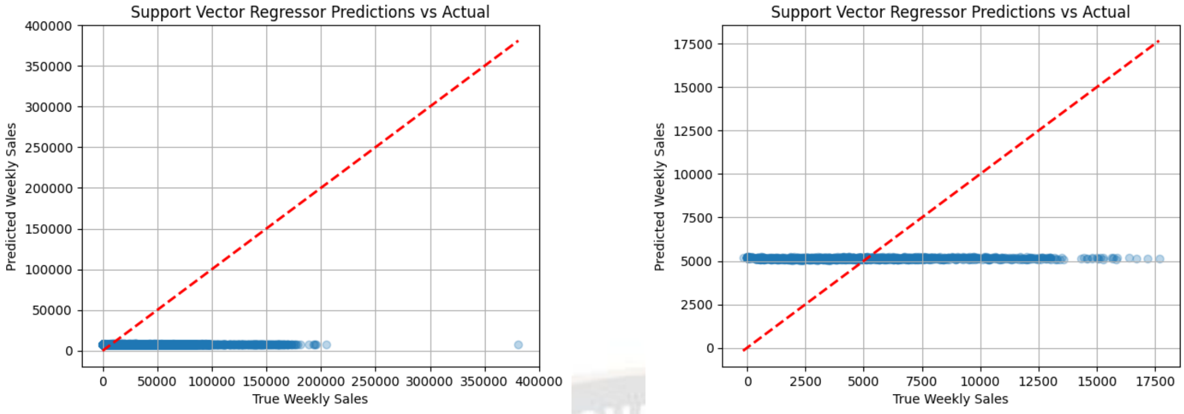
In this section, result evaluation of the customer segmentation based demand forecasting is presented. Customer segmentation is the practice of dividing a company's customers into groups that reflect similarity among customers based on specific characteristics and behaviors. The benefits of segmentation are manifold, from designing tailored marketing campaigns to optimizing product development strategies. In the realm of sales forecasting, the use of customer segmentation could potentially provide richer insights and enhance prediction accuracy. Below Fig 3 to fig 9 shows the result of machine learning (RFR, SVR, LR, kNNR, GBR, DTR, XGBR) with and without customer segmentation.



(a) Without Customer Segmentation (b) With Customer Segmentation

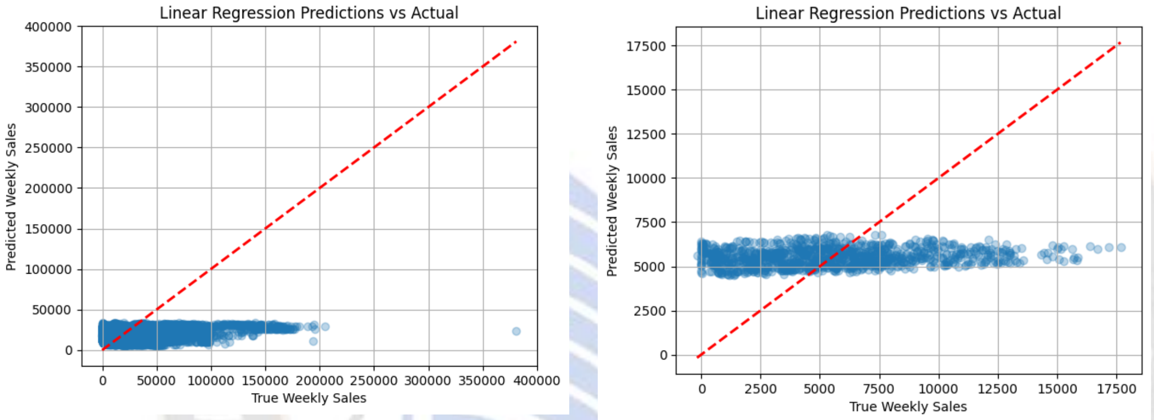
Figure 3: Sales Forecasting with and Without Customer Segmentation using Random Forest Regressor

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 11s
DOI: https://doi.org/10.17762/ijritcc.v11i11s.8170
***Article Received:** 24 June 2023 **Revised:** 24 August 2023 **Accepted:** 09 September 2023*



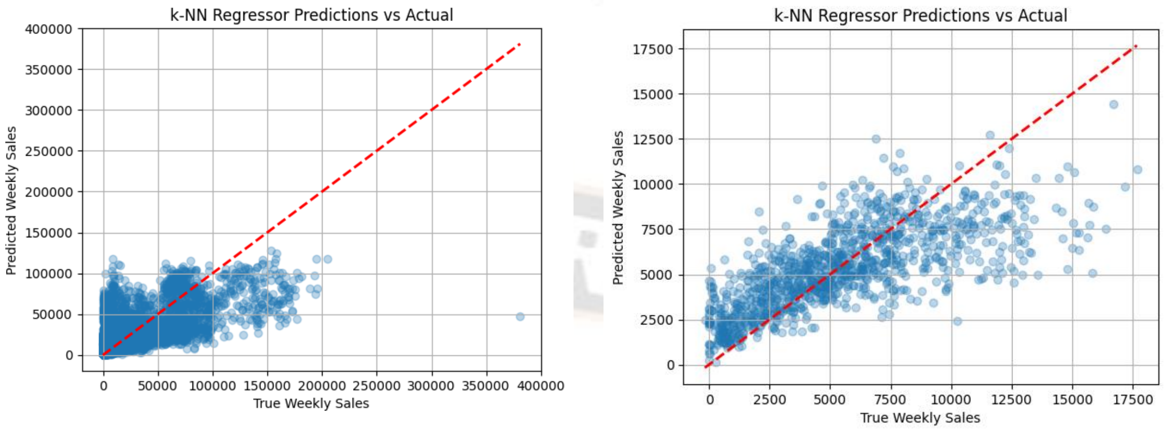
(a) Without Customer Segmentation (b) With Customer Segmentation

Figure 4: Sales Forecasting with and Without Customer Segmentation using Support Vector Regressor



(a) Without Customer Segmentation (b) With Customer Segmentation

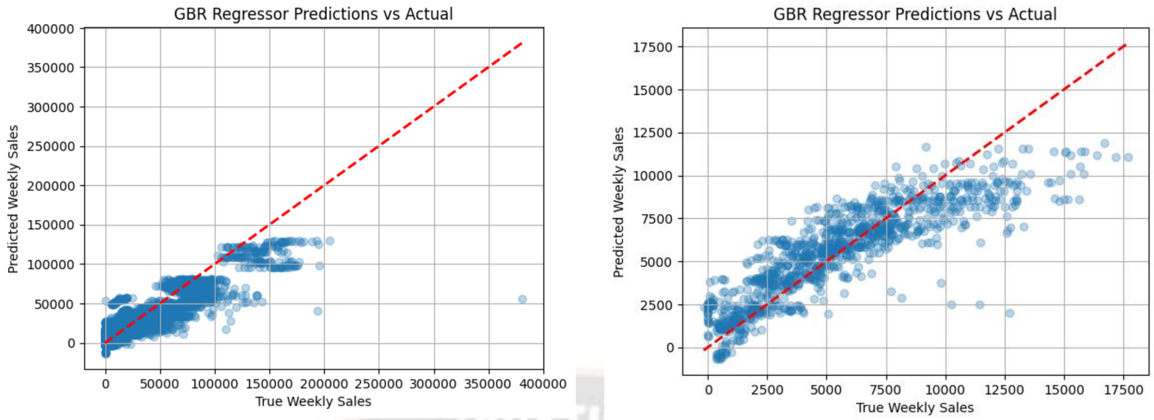
Figure 5: Sales Forecasting with and Without Customer Segmentation using Linear Regressor



(a) Without Customer Segmentation (b) With Customer Segmentation

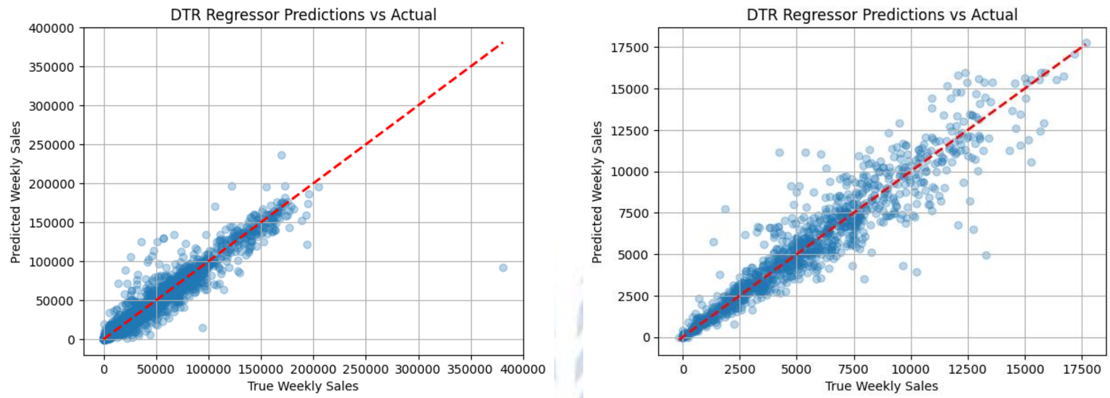
Figure 6: Sales Forecasting with and Without Customer Segmentation using k-NN Regressor

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 11s
DOI: https://doi.org/10.17762/ijritcc.v11i11s.8170
***Article Received:** 24 June 2023 **Revised:** 24 August 2023 **Accepted:** 09 September 2023*



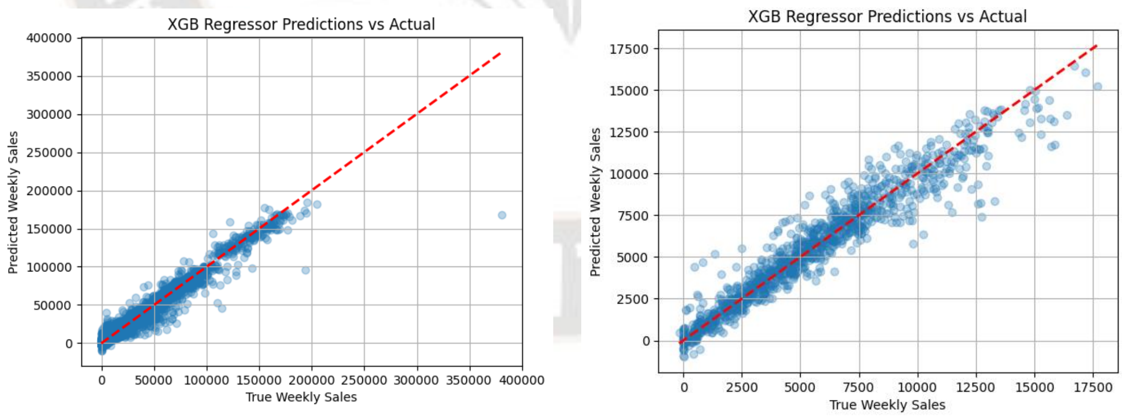
(a) Without Customer Segmentation (b) With Customer Segmentation

Figure 7: Sales Forecasting with and Without Customer Segmentation using Gradiant Boosting Regressor



(a) Without Customer Segmentation (b) With Customer Segmentation

Figure 8: Sales Forecasting with and Without Customer Segmentation using Decision Tree Regressor



(a) Without Customer Segmentation (b) With Customer Segmentation

Figure 9: Sales Forecasting with and Without Customer Segmentation using Extreme Gradient Boosting Regressor

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 11s
DOI: https://doi.org/10.17762/ijritcc.v11i11s.8170
***Article Received:** 24 June 2023 **Revised:** 24 August 2023 **Accepted:** 09 September 2023*

Table 1: Performance Evaluation of Sales Forecasting with and Without Customer Segmentation

Models	Without Customer Segmentation		With Customer Segmentation	
	MAE	**R^2**	**MAE**	**R^2**
Random Forest Regressor	1838.19	0.96	941.77	0.96
Support Vector Regressor	13902.9	-0.13	5956	0
Linear Regressor	16326.71	0.05	5990.68	0.13
k-NN Regressor	10691.94	0.48	3920.43	0.58
Decision Tree Regressor	7382.48	0.76	3692.73	0.64
Gradiant Boosting Regressor	2208.7	0.92	1093.70	0.94
Extreme Gradiant Boosting Regressor	2559.19	0.96	1739.19	0.91
TL-ML	1681.059	0.96	953.77	0.965

The table 1 provides a performance evaluation of various machine learning models when applied to sales forecasting tasks. It compares the performance of these models with and without customer segmentation. Two metrics are used to assess model performance: MAE and R^2.

Effect of Customer Segmentation: In general, incorporating customer segmentation significantly improves the performance (reduces the MAE) of all models when compared to not using customer segmentation.

Best Performing Model: When not using customer segmentation, the 'Random Forest Regressor performs the best models, as evidenced by their low MAE and high R2 with customer segmentation.

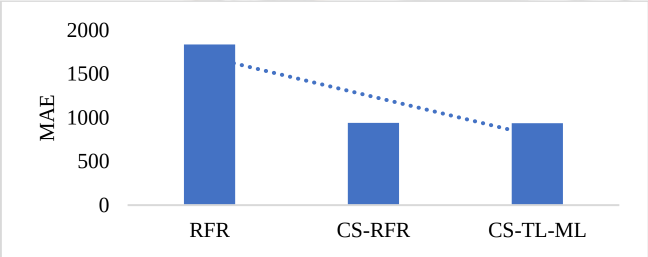


Figure 10: Comparison of MAE for ML Approaches for Demand Forecasting

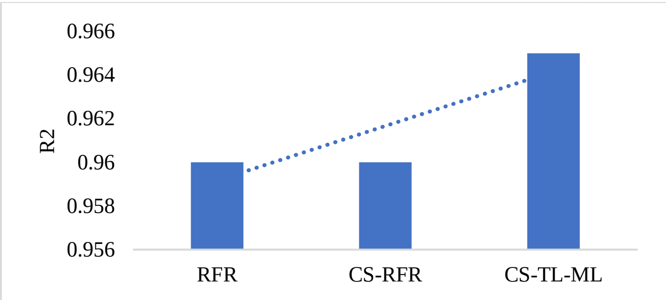


Figure 11: Comparison of R2 for ML Approaches for Demand Forecasting

The fig 10 presents the performance of various techniques, as measured by the Mean Absolute Error (MAE). Lower MAE values indicate better model performance. In summary, the table compares the performance of three techniques, with the ones incorporating customer segmentation (CS-RFR and CS-TL-ML) outperforming the standard Random Forest Regressor. Among the customer segmentation techniques, CS-TL-ML slightly edges out CS-RFR in terms of prediction accuracy. The fig 11 presents the performance of various techniques, as measured by the R2. High R2 values indicate better model performance. In summary, the table compares the performance of three techniques, with the ones incorporating customer segmentation (CS-RFR and CS-TL-ML) outperforming the standard Random Forest Regressor. Among the customer segmentation techniques, CS-TL-ML slightly edges out CS-RFR in terms of prediction accuracy.

V. Conclusion

Our research utilized a hybrid approach combining the BIRCH algorithm and time-lagged machine learning to improve sales forecasting and customer segmentation on a Walmart dataset. The performance metrics used for evaluation were MAE and R2. The results confirm the efficacy of incorporating customer segmentation in sales prediction tasks, as demonstrated by reduced MAE and increased R2 scores for all models examined. Among the models, the ones incorporating customer segmentation (CS-RFR and CS-TL-ML) markedly outperformed the standard Random Forest Regressor. In a head-to-head comparison between the segmentation models, CS-TL-ML emerged with slightly better performance metrics. These findings underscore the importance of customer segmentation in sales prediction and suggest that our hybrid approach could serve as a robust framework for business development, especially for enterprises with large-scale datasets. Future work could extend this methodology to other business domains and include additional features for an even more holistic business development strategy.

References

[1] C. Shao, Y. Yang, S. Juneja, and T. GSeetharam, "IoT data visualization for business intelligence in corporate finance," Inf. Process. Manag., vol. 59, no. 1, p. 102736, Jan. 2022, doi: 10.1016/J.IPM.2021.102736.

[2] A. Gehlot, B. K. Ansari, D. Arora, H. Anandaram, B. Singh, and J. L. Arias-Gonzáles, "Application of Neural Network in the Prediction Models of Machine Learning Based Design," Proc. 2022 Int. Conf. Innov. Comput. Intell. Commun. Smart Electr. Syst. ICSES 2022, 2022, doi: 10.1109/ICSES55317.2022.9914184.

[3] Ali, Naeem, Taher M. Ghazal, Alia Ahmed, Sagheer Abbas, M. A. Khan, Haitham M. Alzoubi, Umar Farooq, Munir Ahmad, and Muhammad Adnan Khan. "Fusion-based supply chain collaboration using machine learning techniques." Intelligent Automation and Soft Computing 31, no. 3 (2022): 1671-1687.

[4] M. Huber, J. Meier, and H. Wallimann, "Business analytics meets artificial intelligence: Assessing the demand effects of discounts on Swiss train tickets," Transp. Res. Part B Methodol., vol. 163, pp. 22–39, Sep. 2022, doi: 10.1016/J.TRB.2022.06.006.

[5] D. Irfan, X. Tang, V. Narayan, P. K. Mall, S. Srivastava, and V. Saravanan, "Prediction of Quality Food Sale in Mart Using the AI-Based TOR Method," J. Food Qual., vol. 2022, 2022, doi: 10.1155/2022/6877520.

[6] M. Hawkins, "Metaverse Live Shopping Analytics: Retail Data Measurement Tools, Computer Vision and Deep Learning Algorithms, and Decision Intelligence and Modeling," J. Self-Governance Manag. Econ., vol. 10, no. 2, pp. 2377–0996, 2022, doi: 10.22381/jsme10220222.

[7] R. Patriarca, G. Di Gravio, R. Cioponea, and A. Licu, "Democratizing business intelligence and machine learning for air traffic management safety," Saf. Sci., vol. 146, p. 105530, Feb. 2022, doi: 10.1016/J.SSCI.2021.105530.

[8] M. N. Omri and W. Mribah, "Intelligent Systems and Applications, 2022, 1, 1-23J. Intelligent Systems and Applications," vol. 1, pp. 1–23, 2022, doi: 10.5815/ijisa.2022.01.01.

[9] Watson, Robert. "The virtual economy of the metaverse: Computer vision and deep learning algorithms, customer engagement tools, and behavioral predictive analytics." Linguistic and Philosophical Investigations 21 (2022): 41-56.

[10] A. Shewale, A. Mokhade, N. Funde, and N. D. Bokde, "A Survey of Efficient Demand-Side Management Techniques for the Residential Appliance Scheduling Problem in Smart Homes," Energies 2022, Vol. 15, Page 2863, vol. 15, no. 8, p. 2863, Apr. 2022, doi: 10.3390/EN15082863.

[11] Y. Niu, L. Ying, J. Yang, M. Bao, and C. B. Sivaparthipan, "Organizational business intelligence and decision making using big data analytics," Inf. Process. Manag., vol. 58, no. 6, p. 102725, Nov. 2021, doi: 10.1016/J.IPM.2021.102725.

[12] M. D. Tamang, V. Kumar Shukla, S. Anwar, and R. Punhani, "Improving Business Intelligence through Machine Learning Algorithms," Proc. 2021 2nd Int. Conf. Intell. Eng. Manag. ICIEM 2021, pp. 63–68, Apr. 2021, doi: 10.1109/ICIEM51511.2021.9445344.

[13] C. A. Tavera Romero, J. H. Ortiz, O. I. Khalaf, and A. R. Prado, "Business Intelligence: Business Evolution after Industry 4.0," Sustain. 2021, Vol. 13, Page 10026, vol. 13, no. 18, p. 10026, Sep. 2021, doi: 10.3390/SU131810026.

[14] J. Ranjan and C. Foropon, "Big Data Analytics in Building the Competitive Intelligence of Organizations," Int. J. Inf. Manage., vol. 56, p. 102231, Feb. 2021, doi: 10.1016/J.IJINFOMGT.2020.102231.

[15] Sakura Nakamura, Machine Learning in Environmental Monitoring and Pollution Control , Machine Learning Applications Conference Proceedings, Vol 3 2023.

[16] A. J. Nakhal A, R. Patriarca, G. Di Gravio, G. Antonioni, and N. Paltrinieri, "Investigating occupational and operational industrial safety data through Business Intelligence and Machine Learning," J. Loss Prev. Process Ind., vol. 73, p. 104608, Nov. 2021, doi: 10.1016/J.JLP.2021.104608.

[17] J. Z. Zhang, P. R. Srivastava, D. Sharma, and P. Eachempati, "Big data analytics and machine learning: A retrospective overview and bibliometric analysis," Expert Syst. Appl., vol. 184, p. 115561, Dec. 2021, doi: 10.1016/J.ESWA.2021.115561.

[18] K. Annapurani, E. Poovammal, C. Ruvinga, and I. Venkat, "Healthcare Data Analytics Using Business Intelligence Tool," Mach. Learn. Anal. Healthc. Syst., pp. 191–212, Jun. 2021, doi: 10.1201/9781003185246-10.

[19] Prof. Arun Pawar, Mr. Dharmesh Dhabliya. (2018). Intelligent Modulation Recognition System and its Implementation using MATLAB. International Journal of New Practices in Management and Engineering, 7(01), 08 - 14. https://doi.org/10.17762/ijnpme.v7i01.63.

[20] S. Fraihat, W. A. Salameh, A. Elhassan, B. A. Tahoun, and M. Asasfeh, "Business Intelligence Framework Design and Implementation: A Real-estate Market Case Study," ACM J. Data Inf. Qual., vol. 13, no. 2, Jun. 2021, doi: 10.1145/3422669.

[21] H. W. Khan et al., "Intelligent Optimization Framework for Efficient Demand-Side Management in Renewable Energy Integrated Smart Grid," IEEE Access, vol. 9, pp. 124235–124252, 2021, doi: 10.1109/ACCESS.2021.3109136.

[22] A. J. Silva, P. Cortez, C. Pereira, and A. Pilastri, "Business analytics in Industry 4.0: A systematic review," Expert Syst., vol. 38, no. 7, p. e12741, Nov. 2021, doi: 10.1111/EXSY.12741.

[23] https://www.kaggle.com/competitions/walmart-recruiting-store-sales-forecasting/data