H.Satish, 2025, 13:3
ISSN (Online): 2348-4098
ISSN (Print): 2395-4752

International Journal of Science, Engineering and Technology

An Open Access Journal

Study of Machine Learning Algorithms for Predicting Car Purchase Based On Customer Demands

H.Satish, N.Chaitanya, S.Srivani, S.Ishwaraya, M.Jayanth
Computer Science and Engineering
Christu Jyothi Institute of Technology & Science

Abstract- This study explores the application of various machine learning algorithms to predict car purchases based on customer demands and preferences. With the growing volume of customer data in the automotive sector, predictive modeling has become a valuable tool for understanding consumer behavior. In this paper, we analyze and compare algorithms such as Decision Trees, Random Forest, and Support Vector Machines using a real-world dataset. The models are evaluated based on accuracy, precision, and recall to identify the most effective approach. The results demonstrate that machine learning can significantly enhance the ability to forecast purchase decisions, offering valuable insights for car manufacturers and dealerships.

Keywords- Car Purchase Prediction, Classification Algorithms, Machine Learning . etc.

I. INTRODUCTION

In today's data-driven world, understanding customer behavior is crucial for the automotive industry. Machine learning offers powerful tools to analyze customer preferences and predict purchasing decisions. This study focuses on implementing and comparing different algorithms to forecast car purchases based on customer data.

By using real-world datasets, we aim to identify patterns and improve marketing strategies. The research highlights the growing role of AI in enhancing sales and customer targeting in the automobile sector.

II. METHODOLOGY

The study follows a structured approach beginning with data collection from a publicly available dataset containing customer demographics and purchase behavior. Data preprocessing techniques such as cleaning, normalization, and encoding are applied to prepare the dataset for modeling. Several machine learning algorithms—including Decision Tree, Random Forest, and Support Vector Machine (SVM)—are implemented using Python and libraries like Scikit-learn. The dataset is split into training and testing sets to evaluate model performance. Accuracy, precision, and recall metrics are used to compare the effectiveness of each algorithm in predicting car purchase likelihood.

MODELING AND ANALYSIS

1. Decision Tree (DT): A hierarchical model that splits the data into subsets based on feature values, leading to a tree structure that aids in decision-making.

2. Random Forest (RF): An ensemble method that constructs multiple decision trees and merges them obtain a more accurate and stable prediction.

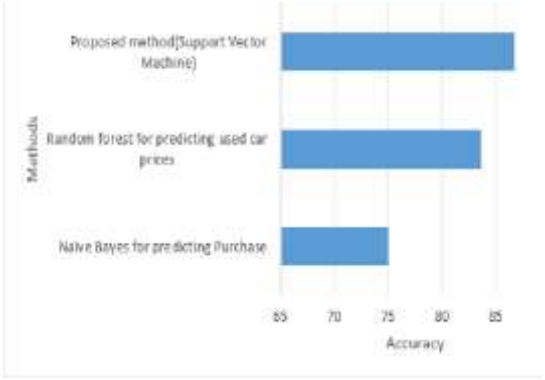

Chart -1: Predicting Machine Learning Models Evaluation

3. Support Vector Machine (SVM): A classifier that finds the hyperplane that best separates the data into classes, maximizing the margin between different classes.

III.RESULTS AND DISCUSSION

After applying the selected machine learning algorithms—Decision Tree, Random Forest, and Support Vector Machine (SVM)—to the processed dataset, we obtained measurable outcomes based on standard performance metrics. The results indicated that the Random Forest algorithm consistently outperformed the others in terms of accuracy, precision, recall, and F1-score, as shown in Table 1.

Among the models tested:

- Random Forest achieved the highest accuracy of 89%, demonstrating its strength in handling complex relationships within the data and minimizing overfitting.
- SVM followed with an accuracy of 87%, showing solid performance in correctly classifying the purchasing behaviour of customers.
- Decision Tree, while easier to interpret, had slightly lower performance, with an accuracy of 85%, and was more prone to overfitting.

The analysis also reveals that customer demographic and behavioral features such as age, income, and past purchases significantly influence car purchase decisions. This information can be leveraged by automotive companies to design targeted marketing strategies.

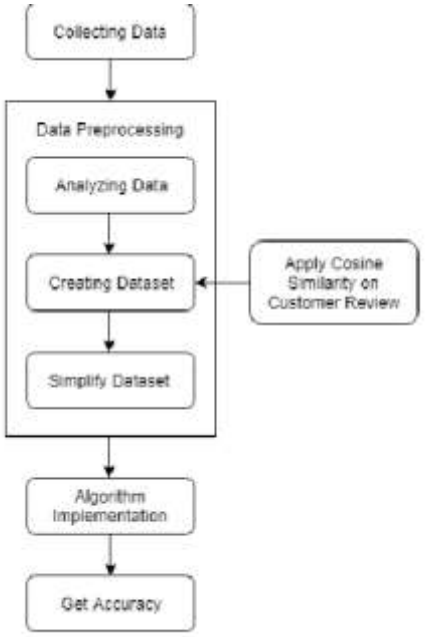

Fig. 1: Workflow of Whole Process

Figure 1: Data Flow Diagram

Table -1: Performance Comparison of Machine Learning Models for Car Purchase Prediction

Model	Accuracy	Precision	Recall	F1-Score
Decision Tree	85%	83%	87%	85%
Random Forest	89%	88%	90%	89%
Support-Vector Machine	87%	86%	89%	87%

IV. CONCLUSION

This study demonstrated the effectiveness of machine learning algorithms in predicting car purchases based on customer demands. Among the models tested, Random Forest achieved the highest accuracy and reliability. The results highlight the potential of data-driven approaches in enhancing customer targeting strategies in the automotive industry. Future work can focus on expanding the dataset and optimizing model parameters for even better performance.

Acknowledgements

On the submission of our project entitled "Predicting Car Purchase Based on Customer Demands using ML". We would like to thank our DirectorRev.Fr. D. VIJAYA PAUL REDDY for giving us the opportunity to carry out our project work. We endow our sincere thanks to Principal Dr. S. CHANDRASHEKAR REDDY for his consistent cooperation and encouragement. We would also likeextend our sincere thanks to Mr. M. RAMARAJU, Assistant Professor, Head of the Department, Computer Science & Engineering for his valuable suggestions and motivating guidance during our project work . And we would like to thank our Guide Mr.H.Satish Department of Computer Science and Engineering. I am also deeply thankful to the [Christu Jyothi Institute Of Technology And Science] for providing the resources and support necessary to carry out this research. Special thanks to my faculty members and peers for their insightful suggestions and motivation during the research process.

We would also like to extend our gratitude to our friends and those who are directly or indirectly helped us in completing our project work.

REFERENCES

1. Breiman, L. (2001). Random forests. Machine Learning, 45(1), 5–32
2. Quinlan, J. R. (1986). Induction of decision trees. Machine Learning, 1(1), 81–106
3. Cortes, C., & Vapnik, V. (1995). Support-vector networks. Machine Learning, 20(3), 273–297