IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

Integrating Supervised Machine Learning with Real-Time 3D Simulation for Customer Purchase Behavior Modeling

1st Author : Raydah mohammed Al-Ajmi
2nd Author : Dr. Ahmad M.Alkheder Almasabi

Abstract

This project presents a supervised machine learning approach for predicting customer purchase behavior and integrating the trained model into a Blender-based simulation environment. The dataset includes age and estimated salary as input variables and purchase decision as the target variable. Three classification models—Logistic Regression, K-Nearest Neighbors (KNN), and Random Forest—were implemented and evaluated. Performance was measured using accuracy, precision, recall, F1-score, and confusion matrix analysis. The KNN model achieved the highest accuracy and was exported for integration into Blender simulation.

I. Introduction

Modeling and simulation play an essential role in representing real-world systems and supporting decision-making processes. With the rapid growth of data-driven technologies, machine learning techniques have become widely used for predictive analysis in various domains, including marketing and customer behavior prediction [1], [5].

Customer purchase prediction is a classification problem in which demographic attributes are used to determine whether a customer will buy a product. Supervised learning algorithms are particularly suitable for such tasks, as they learn patterns from labeled datasets and generate predictions for unseen data.

In this project, three supervised machine learning models—Logistic Regression [4], K-Nearest Neighbors (KNN) [2], and Random Forest [3]—are implemented to predict customer purchasing behavior based on Age and Estimated Salary. The models are evaluated using standard performance metrics, including accuracy, precision, recall, F1-score, and confusion matrix analysis.

The best-performing model is then exported and integrated into a Blender simulation environment. This integration demonstrates how machine learning prediction can be combined with 3D simulation to model customer behavior dynamically.

The objective of this project is to develop an accurate purchase prediction model and demonstrate its practical application within a simulation framework.

II. Dataset Description

The dataset used is the Social Network Ads dataset.

Selected input features:

- Age
- Estimated Salary

Target variable:

- Purchased (0 = No, 1 = Yes)

The dataset was divided into:

- 75% training data
- 25% testing data

III. Methodology

A. Data Preprocessing

The following preprocessing steps were applied:

1. Feature selection (Age and EstimatedSalary)
2. Train-test split (test size = 0.25)
3. Feature scaling using StandardScaler

Feature scaling is particularly important for distance-based algorithms such as KNN [2].

B. Classification Models

Three supervised classification models were implemented:

1. Logistic Regression [4]
2. K-Nearest Neighbors (k = 5) [2]
3. Random Forest (100 estimators) [3]

Each model was trained using the training dataset and evaluated on the testing dataset.

IV. Evaluation Metrics

Model performance was evaluated using:

- Accuracy
- Precision
- Recall
- F1-score

- Confusion Matrix

These metrics provide a comprehensive assessment of classification performance [5].

V. Experimental Results

A. Model Performance Comparison

Model	Accuracy	Precision	Recall	F1 Score
Logistic Regression	0.89	0.8889	0.75	0.8136
KNN	0.93	0.8788	0.9063	0.8923
Random Forest	0.92	0.8529	0.9063	0.8788

The KNN model achieved the highest accuracy of 93%.

B. Confusion Matrix (KNN)

The confusion matrix for the KNN model indicates:

- 64 true negative predictions
- 29 true positive predictions
- 4 false positive predictions
- 3 false negative predictions

This confirms the strong predictive capability of the KNN classifier.

VI. Integration with Blender Simulation

After training, the best-performing model (KNN) and the fitted StandardScaler were exported using joblib:

- best_model.pkl
- scaler.pkl

Within Blender:

1. Customer demographic attributes (Age and Estimated Salary) are provided as input.
2. The scaler transforms the input features.
3. The trained KNN model generates a prediction.
4. The simulation responds according to the predicted output.

If the prediction equals 1, the simulated customer performs a purchase action.
If the prediction equals 0, the customer does not purchase.

This demonstrates practical integration between machine learning prediction and 3D simulation modeling.

VII. Conclusion

This project demonstrated the use of supervised machine learning techniques for predicting customer purchasing behavior and integrating the trained model into a Blender simulation environment.

Among the tested models, KNN achieved the highest performance with 93% accuracy. The exported model enables simulation-driven behavioral modeling.

Acknowledgment

The author would like to thank Dr. Ahmed Mohammed Nasser Al-Khader Al-Masaabi for guidance during the Modeling and Simulation course.

References

[1] C. M. Bishop, *Pattern Recognition and Machine Learning*. Springer, 2006.

[2] T. Cover and P. Hart, "Nearest neighbor pattern classification," IEEE Transactions on Information Theory, vol. 13, no. 1, pp. 21–27, 1967.

[3] L. Breiman, "Random forests," Machine Learning, vol. 45, no. 1, pp. 5–32, 2001.

[4] D. W. Hosmer, S. Lemeshow, and R. X. Sturdivant, *Applied Logistic Regression*, 3rd ed., Wiley, 2013.

[5] J. Han, M. Kamber, and J. Pei, *Data Mining: Concepts and Techniques*, 3rd ed., Morgan Kaufmann, 2011.