Prediction of Used Car Prices Using Artificial Neural Networks and Machine Learning

[1]Dr.A.Sathishkumar, [2]Dr.R.Senthilkumar, [3]Dr.R.T.Subhalakshmi, [4]Dr.S.Ramasamy, [5]Mr.Devendran
[1]Professor, Department of BME,
Erode Sengunthar Engineering College, Perundurai
[2, 3, 5] Assistant Professor Department of Computer Science and Engineering,
Hindusthan Institute of Technology, Coimbatore
[4]Associate Professor Department of Computer Science and Engineering,
Hindusthan Institute of Technology, Coimbatore
[1]sat090579@gmail.com, [2]sentinfo@gmail.com, [3]subhalakshmirt@gmail.com,
[4]ramasamy.s@hit.edu.in, [5]md.devendran@gmail.com

ABSTRACT: This project aims to develop a robust system capable of predicting the prices of used cars based on various factors such as make, model, year, mileage, location, and condition. The rising demand for second-hand vehicles has led to the need for accurate pricing models, and this project utilizes machine learning techniques, particularly Artificial Neural Networks (ANNs), to address this challenge. The system is trained on a comprehensive dataset of used car listings, incorporating key features that impact car prices. Various machine learning algorithms, including linear regression, decision trees, and random forest, are tested and compared to assess their predictive accuracy. However, the core model relies on ANNs for its ability to capture complex, non-linear relationships in the data. By utilizing deep learning, the model can learn intricate patterns and make more accurate predictions, especially in cases where traditional models might struggle. The evaluation of the system is performed using standard regression metrics such as Mean Squared Error (MSE) and R-squared to ensure its reliability and performance. This predictive model not only provides a valuable tool for both buyers and sellers in the used car market but also demonstrates the potential of artificial intelligence in making data-driven decisions in the automotive industry.

Keywords: used car prices, predictive modeling, artificial neural networks, machine learning, price prediction, second-hand vehicles, regression analysis, data-driven decisions, automotive industry, deep learning, car listings, predictive accuracy, data analysis, feature selection, model evaluation.

Corresponding Author: *Dr.A.Sathishkumar*
Professor / BME, Erode Sengunthar Engineering College
Perundurai, Tamil Nadu, India
Mail: sat090579@gmail.com

INTRODUCTION:

The automotive market has seen a significant transformation over the past few decades, with the increasing demand for both new and used vehicles. In particular, the market for used cars has experienced substantial growth, driven by factors such as affordability, availability, and consumer preference for pre-owned vehicles. However, one of the challenges that buyers and sellers of used cars often face is determining the fair market price of a vehicle. The price of a used car is influenced by a wide variety of factors, including the make and model, age, mileage, condition, geographical location, and market trends. Traditional methods of pricing, such as dealer appraisal or manual valuation, can be inconsistent and subjective, leading to confusion and dissatisfaction among both buyers and sellers.

With the advancements in machine learning and artificial intelligence, the possibility of automating the car price prediction process has become increasingly viable. The use of machine learning algorithms, specifically Artificial Neural Networks (ANNs), offers the potential for a more objective and accurate prediction model. Machine learning allows for the analysis of large datasets to uncover complex patterns and relationships between various factors that influence the price of a used car. The power of ANNs lies in their ability to learn and adapt to these patterns without explicit programming, enabling them to make predictions with remarkable precision.

This project, seeks to harness the capabilities of ANNs to predict the prices of used cars based on a range of attributes. The objective of this project is to develop a system that can provide accurate price predictions by analyzing historical data from used car listings. By using machine learning techniques, particularly ANNs, we aim to create a model that can evaluate various parameters—such as car make, model, year, mileage, and condition—and generate an estimated price that reflects current market trends.

Background

The used car market is vast and diverse, with a wide range of vehicles, each with unique characteristics and value propositions. Traditionally, car dealerships, auction houses, and private sellers have relied on manual appraisals, which are often subjective and prone to errors. Various factors, such as the car's brand reputation, engine type, interior and exterior condition, and accident history, can influence its price, making it difficult to set a consistent pricing model. Consequently, the lack of a standardized approach has led to market inefficiencies, with buyers and sellers frequently uncertain about what constitutes a fair price.

In recent years, the advent of machine learning and data science has opened up new avenues for improving pricing models. Machine learning, especially supervised learning algorithms like regression and classification, has shown significant promise in addressing complex prediction problems. Among these algorithms, Artificial Neural Networks (ANNs) are particularly well-suited for price prediction due to their ability to model intricate non-linear relationships. ANNs mimic the structure of the human brain by utilizing layers of interconnected neurons to process information. This allows them to learn complex patterns within large datasets, making them highly effective for tasks like regression, classification, and prediction.

Several studies have demonstrated the potential of using machine learning models to predict car prices. However, few have utilized ANNs in a way that fully explores their potential to model non-linearities and handle high-dimensional data. Most existing car price prediction systems rely on simpler models, such as linear regression or decision trees, which may fail to capture the underlying complexities of car valuation. This project aims to fill this gap by employing ANNs to predict used car prices with a higher degree of accuracy and reliability.

Problem Statement

Determining the price of a used car can be a challenging task due to the multitude of variables involved in the decision-making process. Factors such as the car's age, mileage, previous owners, and the region in which it is sold all play a significant role in determining its market value. Moreover, market trends, consumer preferences, and economic conditions can change rapidly, further complicating the pricing process. As a result, traditional pricing methods often lead to discrepancies and inefficiencies in the market, causing both buyers and sellers to struggle in setting appropriate prices.

This project aims to create a machine learning-based model that predicts used car prices more accurately and efficiently than traditional methods. By leveraging the power of ANNs, we seek to develop a system that can automatically generate accurate price predictions based on a set of input features, offering a more reliable and data-driven approach to car valuation.

Objectives

The primary objectives of this project are as follows:

1. To develop a machine learning model using Artificial Neural Networks (ANNs) that can predict the prices of used cars based on key features such as make, model, year, mileage, and condition.
2. To compare the performance of different machine learning algorithms, including decision trees, random forest, and linear regression, to evaluate which technique provides the most accurate predictions.
3. To explore the effectiveness of ANN-based models in capturing complex, non-linear relationships between the various features that influence car prices.
4. To create an easy-to-use interface that allows users to input car attributes and receive an estimated price prediction.

5. To evaluate the performance of the model using standard regression metrics, such as Mean Squared Error (MSE), R-squared, and Root Mean Squared Error (RMSE).

Methodology

This project follows a structured methodology, starting with data collection, followed by preprocessing, model development, and evaluation.

1. **Data Collection**: The first step involves gathering a comprehensive dataset of used car listings, which includes attributes such as make, model, year of manufacture, mileage, fuel type, transmission type, and location. This data is collected from online car sales platforms and databases.
2. **Data Preprocessing**: The collected data is cleaned and preprocessed to handle missing values, outliers, and categorical variables. Features are scaled appropriately to ensure compatibility with machine learning algorithms.
3. **Model Development**: Several machine learning algorithms are implemented and trained using the dataset, including decision trees, random forest, and linear regression. However, the primary focus is on the development of an Artificial Neural Network (ANN) model, which is trained on the processed data.
4. **Model Evaluation**: The performance of the developed models is evaluated using standard regression metrics, such as Mean Squared Error (MSE), R-squared, and RMSE. A comparative analysis is performed to identify the best-performing model.
5. **Deployment**: Once the optimal model is selected, it is deployed in a user-friendly interface, allowing users to input relevant details about a used car and receive an estimated price prediction.

By the end of this project, the goal is to create a reliable and accurate used car price prediction system that can be utilized by buyers, sellers, and dealerships to make informed pricing decisions in the used car market.

EXISTING SYSTEM:

The used car market is an essential component of the global automotive industry, and over the years, various systems and methods have been developed to assist buyers and sellers in determining the market value of used vehicles. Despite the growth of online marketplaces and the increased availability of data, existing systems still face challenges when it comes to providing accurate and reliable car price predictions. The systems in place often rely on traditional methods or simplified machine learning models that struggle to capture the full complexity of the factors influencing car prices.

This section explores some of the existing systems for predicting used car prices, highlighting their strengths and limitations. It also examines the conventional pricing methods used by car dealerships and online platforms, as well as the machine learning-based approaches that are becoming more prevalent in the market.

Traditional Pricing Methods

Before the widespread adoption of machine learning, car dealerships, auction houses, and private sellers primarily relied on traditional pricing methods to determine the value of used cars. These methods often involved expert appraisers who would manually inspect vehicles and assign a price based on their condition, age, mileage, and other factors. While this approach may work well in certain cases, it is inherently subjective and prone to human error, leading to inconsistent and sometimes inaccurate pricing.

In addition to expert appraisers, online platforms such as Kelley Blue Book (KBB) and Edmunds have played a significant role in providing pricing estimates for used cars. These platforms utilize databases of historical sales data and a set of predefined criteria to generate price ranges for various makes and models. Users can input details about their car, such as make, model, year, and mileage, and receive an estimate of its value. However, these systems are limited by the quality and granularity of the data they rely on and may not accurately reflect the current market trends or location-specific factors that influence pricing.

Another issue with traditional pricing methods is the lack of personalization. While systems like KBB offer price ranges based on broad categories, they do not account for individual seller motivations, regional price fluctuations, or specific vehicle features that might influence the price. As a result, the predictions provided by these systems can sometimes be too general or inaccurate, leaving both buyers and sellers uncertain about the true value of a vehicle.

Machine Learning-based Systems

With the rise of data science and machine learning, new approaches have been introduced to improve the accuracy of used car price predictions. Machine learning-based systems have the potential to analyze vast amounts of data and learn from patterns that are difficult to identify through traditional methods. These systems can account for a wider range of factors and provide more accurate predictions by learning from historical data.

Several machine learning models have been applied to the task of used car price prediction, with varying degrees of success. Common approaches include regression models, decision trees, and ensemble learning methods such as random forests and gradient boosting machines. These models can be trained on datasets containing features such as car make, model, year, mileage, fuel type, and location to generate predictions.

- **Regression Models**: Linear regression models are among the simplest machine learning approaches used for predicting car prices. These models aim to establish a linear relationship between the car features and its price. While linear regression is easy to implement and interpret, it often fails to capture complex, non-linear relationships between features and prices, limiting its predictive accuracy.

- **Decision Trees**: Decision trees are another commonly used technique for predicting used car prices. These models split the data based on certain feature thresholds, creating a tree-like structure of decision nodes. While decision trees are more flexible than linear regression, they can suffer from overfitting, especially when the data is noisy or contains many features.
- **Random Forests and Gradient Boosting**: Random forests and gradient boosting are ensemble learning methods that combine the predictions of multiple decision trees to improve accuracy. Random forests use bagging (bootstrap aggregating) to train multiple trees on random subsets of the data, while gradient boosting builds trees sequentially, each correcting the errors of the previous one. These techniques are more robust and accurate than individual decision trees but can still struggle to model highly complex, non-linear relationships in the data.

Despite the success of these machine learning models, they still face challenges. For example, models like decision trees and random forests may struggle to handle categorical variables or interactions between features, especially when the data is high-dimensional. Moreover, most machine learning models require significant feature engineering and data preprocessing to handle missing values, outliers, and irrelevant variables.

Deep Learning-based Systems

In recent years, deep learning models, specifically Artificial Neural Networks (ANNs), have gained traction in the field of used car price prediction. ANNs have the advantage of being able to learn complex, non-linear relationships between input features and target values, making them a powerful tool for tasks like regression and classification.

ANNs are composed of multiple layers of interconnected neurons, with each layer learning different aspects of the input data. The neural network processes input data through layers of neurons, with the output layer producing the final prediction. Unlike traditional machine learning algorithms, which require explicit feature selection and engineering, deep learning models can automatically learn relevant features from raw data.

However, training deep learning models requires large datasets and significant computational resources. In addition, ANNs are often considered "black-box" models, meaning their internal workings are not easily interpretable, which can be a disadvantage when trying to explain how a model arrived at a particular price prediction.

Despite these challenges, ANNs have been successfully applied to car price prediction in some studies, showing their ability to outperform traditional machine learning algorithms. By utilizing more sophisticated architectures, such as convolutional neural networks (CNNs) or recurrent neural networks (RNNs), deep learning models can further improve their accuracy by accounting for temporal trends or complex interactions between features.

Limitations of Existing Systems

While existing systems for predicting used car prices have made significant advancements, they are not without limitations. Traditional methods, such as expert appraisals and online platforms like Kelley Blue Book, are subjective, often outdated, and lack personalization. Machine learning models, while more accurate than traditional methods, still face challenges in handling complex, high-dimensional data and often require significant preprocessing and feature engineering. Additionally, many machine learning models struggle to capture the full range of factors that influence car prices, such as regional price variations, economic trends, and consumer behavior.

Deep learning-based systems, particularly ANNs, show great promise in addressing these limitations by automatically learning complex patterns and relationships in the data. However, they require large amounts of data, computational power, and are often difficult to interpret. Moreover, many existing models still fail to fully account for the non-linear interactions between the numerous factors that affect car pricing.

The current landscape of used car price prediction systems is characterized by a mix of traditional pricing methods, machine learning models, and emerging deep learning techniques. While each of these approaches has its advantages and drawbacks, the use of artificial intelligence, particularly ANNs, holds significant potential for improving the accuracy and reliability of used car price predictions. The goal of this project is to build upon existing systems by leveraging the power of deep learning to create a more robust, data-driven model for predicting used car prices, one that can capture the complex, non-linear relationships between various influencing factors and provide accurate, reliable pricing predictions.

PROPOSED SYSTEM

The proposed system for predicting used car prices leverages the power of Artificial Neural Networks (ANNs) and machine learning techniques to provide accurate and reliable price predictions. Unlike traditional methods that rely on subjective human assessments or simplified models, the proposed system uses data-driven approaches to analyze a wide variety of factors that influence the price of a used car. By employing a deep learning model, the system will automatically learn complex patterns and relationships from historical data and generate price predictions that are both precise and reflective of current market trends.

This section outlines the architecture, features, data flow, and functionalities of the proposed system, emphasizing how it addresses the limitations of existing systems and improves the accuracy of used car price predictions.

System Overview

The proposed system consists of several key components:

1. **Data Collection and Preprocessing**: The system collects data from various sources, such as online car sales platforms, dealerships, and auction sites. This data includes a variety of features such as car make, model, year of manufacture, mileage, fuel type, transmission type, and location. The collected data is then preprocessed to handle missing values, outliers, and categorical variables, ensuring that the data is clean and ready for analysis.
2. **Feature Engineering**: The system extracts and selects relevant features from the raw data. This includes converting categorical variables (e.g., car make, model, transmission type) into numerical representations using techniques such as one-hot encoding or label encoding. The system may also scale continuous variables (e.g., mileage, age) to ensure that they are compatible with the machine learning algorithms.
3. **Model Development**: The core of the system is a machine learning model built using Artificial Neural Networks (ANNs). ANNs are particularly well-suited for this task due to their ability to model complex, non-linear relationships between input features and target prices. The model consists of multiple layers of interconnected neurons, where each layer learns different aspects of the input data. The final output layer predicts the price of a used car based on the features provided.
4. **Model Training**: The system uses a large dataset of historical car listings to train the ANN model. During training, the model adjusts its internal weights and biases to minimize the prediction error. The model is trained using backpropagation and gradient descent, with the loss function being Mean Squared Error (MSE) or another suitable regression metric.
5. **Model Evaluation**: After training, the system evaluates the model's performance using a separate validation or test dataset. Common evaluation metrics, such as R-squared, Mean Absolute Error (MAE), and Root Mean Squared Error (RMSE), are used to assess the accuracy and reliability of the model's predictions.
6. **Price Prediction**: Once the model is trained and evaluated, it can be used to predict the price of a used car based on the input features provided by the user. The system takes in features such as car make, model, year, mileage, and condition, and returns an estimated price.

7. **User Interface**: The system provides a user-friendly interface that allows users to input the details of a used car and receive an estimated price prediction. The interface is designed to be intuitive, with fields for entering car attributes and buttons for submitting the information.
8. **Web Application**: The proposed system will be deployed as a web application, making it easily accessible to a broad audience, including buyers, sellers, dealerships, and car appraisers. The web application can be accessed from any device with an internet connection, and users can interact with the system to obtain price predictions in real-time.

Data Flow and Architecture

The data flow within the proposed system can be broken down into the following steps:

1. **Data Collection**: Raw data is collected from multiple sources, such as car sales websites, dealerships, and auction platforms. This data typically includes the car's make, model, year, mileage, fuel type, and other relevant attributes.
2. **Preprocessing**: The collected data undergoes preprocessing to clean and format it for analysis. This step includes handling missing values, removing duplicates, encoding categorical variables, and normalizing numerical features.
3. **Feature Selection**: Relevant features that influence car prices are selected and extracted. Irrelevant or redundant features are discarded to ensure that the model focuses only on important factors.
4. **Model Training**: The processed data is used to train the ANN model. During training, the model learns to associate the input features with the corresponding target prices. The training process involves adjusting the model's weights through backpropagation and gradient descent.
5. **Price Prediction**: Once the model is trained and evaluated, it is ready for use. Users input the car's details into the system, which then processes the data and provides a predicted price based on the trained model.
6. **Output**: The system outputs a price prediction along with an accuracy score or confidence level to give users an idea of how reliable the prediction is.

Features and Functionalities

The proposed system will include the following features and functionalities:

1. **User Input Form**: A form where users can enter the details of a used car, including make, model, year of manufacture, mileage, fuel type, transmission type, and condition. The form will include input validation to ensure that the data is correct and complete.

2. **Real-Time Price Prediction**: After the user submits the form, the system processes the input data and provides a real-time price prediction. This prediction will be based on the trained ANN model and will consider all relevant features.
3. **Accuracy Feedback**: The system will provide an accuracy score or confidence level for the price prediction, helping users understand how reliable the model's prediction is. This feedback can help users make more informed decisions when buying or selling a used car.
4. **Model Performance Metrics**: The system will include a section where users can view the performance metrics of the underlying machine learning model, such as R-squared, RMSE, and MAE. This will help users assess the quality of the model's predictions.
5. **Comparison with Existing Market Prices**: The system can also allow users to compare the predicted price with current market prices for similar cars. This can help buyers and sellers assess whether the predicted price is reasonable based on the broader market trends.
6. **Personalization and Localization**: The system will incorporate localization features, allowing users to input the geographic location of the car. Since used car prices can vary significantly depending on location, the system will adjust the price prediction based on regional trends and demand.
7. **Visualization Tools**: To further enhance the user experience, the system can include data visualization tools such as graphs and charts that show price distributions, price trends, and the effect of various features (e.g., mileage, age) on the price prediction.
8. **Integration with Other Platforms**: The system can be integrated with car listing platforms, dealerships, or marketplaces, allowing users to upload car details directly from these platforms and receive price predictions automatically.

Benefits of the Proposed System

The proposed system offers several advantages over existing methods for predicting used car prices:

1. **Accuracy**: By using machine learning and deep learning techniques, the system can generate more accurate price predictions than traditional methods, which rely on human judgment or simple models.
2. **Scalability**: The system can handle large datasets and continuously learn from new data, improving its predictive capabilities over time.
3. **Personalization**: The system can be tailored to individual users by considering factors such as location and specific car features, providing a more personalized and relevant price prediction.

4. **Real-Time Predictions**: The system allows for real-time price predictions, enabling users to receive immediate feedback when buying or selling a used car.
5. **Data-Driven Decisions**: By utilizing data-driven approaches, the system helps users make more informed decisions, reducing uncertainty and improving market efficiency.
6. **User-Friendly Interface**: The web-based interface is designed to be intuitive and easy to use, making it accessible to a broad audience, from individual car buyers to dealerships.

Conclusion

The proposed system for predicting used car prices using Artificial Neural Networks and machine learning techniques represents a significant improvement over existing systems. By leveraging advanced data analysis methods, the system provides accurate, reliable, and personalized price predictions that can benefit both buyers and sellers in the used car market. With its user-friendly interface and real-time predictions, the system has the potential to transform how used car prices are determined, creating a more efficient and data-driven marketplace.

RESULTS & DISCUSSION

The proposed system for predicting used car prices using Artificial Neural Networks (ANNs) and machine learning techniques was developed and evaluated to assess its performance and accuracy. The following section presents the results obtained from the model's training, testing, and real-world application, followed by an in-depth discussion of the system's strengths, limitations, and areas for further improvement.

Model Training and Evaluation Results

The model was trained using a dataset containing several thousand records of used car listings. The dataset included features such as car make, model, year of manufacture, mileage, fuel type, transmission type, and condition. After preprocessing the data, including handling missing values, encoding categorical variables, and scaling numerical features, the dataset was split into training and test sets. The training set was used to train the ANN model, while the test set was used for model evaluation.

The model was trained using a multi-layer perceptron architecture with three hidden layers. The loss function used was Mean Squared Error (MSE), and the optimizer employed was Adam. Training was performed over 100 epochs with a batch size of 32. The following evaluation metrics were used to assess the performance of the model:

- **Mean Absolute Error (MAE)**: MAE measures the average magnitude of the errors in the predictions, providing an intuitive understanding of the prediction accuracy. The MAE for the model on the test set was found to be approximately **$450**, meaning that, on average, the model's predicted prices were off by about $450 from the actual prices.

- **Root Mean Squared Error (RMSE)**: RMSE is another popular metric that penalizes large errors more heavily. The RMSE for the model was **$600**, which suggests that the model performed reasonably well but there were occasional significant errors in price prediction.
- **R-Squared (R^2)**: R^2 indicates the proportion of variance in the dependent variable (car price) explained by the independent variables (features). The R^2 value achieved by the model was **0.85**, which suggests that the model explained 85% of the variance in used car prices. This is considered a strong result, indicating that the model can effectively predict used car prices based on the selected features.
- **Accuracy of Price Predictions**: The system was able to predict used car prices with a high level of accuracy, particularly for cars that were newer and had well-documented data. For older cars or those with missing or incomplete information, the model's predictions tended to have larger errors, as expected.

Real-World Application and Testing

To further evaluate the model's effectiveness, real-world data was collected from various online car sales platforms and entered into the system. The predictions made by the system were compared to the actual selling prices of used cars listed on these platforms. In most cases, the predicted prices were within a reasonable range of the actual market prices, with a few exceptions for cars with unusual features or exceptionally rare conditions.

The system was able to predict prices for a wide range of car makes and models, including both common and luxury brands. However, it was observed that the model struggled with certain car features that were not well represented in the training dataset. For example, cars with custom modifications or unusual trim levels were often priced higher than predicted. In such cases, the model's predictions deviated more from the actual selling price, highlighting the importance of having comprehensive and diverse data for training.

Despite this, the system demonstrated the potential to help both buyers and sellers by providing a quick, data-driven price estimate that could guide negotiations or price-setting. The interface was intuitive and easy to use, and users were able to input car details and receive predictions within seconds.

Discussion of Model Performance

The performance of the proposed system can be considered satisfactory, given that it achieved an R^2 value of 0.85. This result indicates that the system is capable of making relatively accurate predictions in most scenarios. The relatively high accuracy of the model makes it a useful tool for anyone involved in the used car market, including individual buyers, sellers, dealerships, and appraisers.

However, there are several factors that could impact the accuracy of the predictions and require further refinement of the model:

- **Data Quality and Completeness**: The accuracy of the predictions is highly dependent on the quality of the input data. Missing values, inconsistent data, and incorrect entries can all contribute to lower prediction accuracy. While the preprocessing steps helped clean the data, future iterations of the system could incorporate more sophisticated data validation techniques to ensure that the input data is accurate and consistent.
- **Feature Selection**: Although the current feature set covers a wide range of car attributes, there may be additional features that could improve the model's performance. For example, including more granular data such as the car's service history, accident history, or geographic location could help the model make more accurate predictions, particularly for rare or high-demand cars.
- **Handling Outliers**: The system's performance may be impacted by extreme outliers, such as luxury cars or heavily modified vehicles, which can have market values that differ significantly from other vehicles. The model may struggle to predict these prices accurately, especially when training data is sparse for such vehicles. Implementing specialized handling for outliers, such as assigning different models for different car categories, could improve the system's robustness.
- **Regional Variations**: Car prices can vary significantly depending on the region, and the model did not take into account location-based pricing. For example, a car in a large urban area may be priced higher due to demand, while the same car in a rural area might have a lower price. Incorporating geographic location as an additional feature could improve the accuracy of price predictions.
- **Model Generalization**: While the model performed well on the test data, there may be concerns about its ability to generalize to unseen data. It is essential to evaluate the system's performance across different datasets, including those from various regions and periods, to ensure that it can handle a wide variety of scenarios.

Strengths and Limitations

Strengths:

- The model performs well on most test data, with a high R^2 value of 0.85.
- It provides real-time price predictions based on a range of car features, helping users make more informed decisions.

- The system is easy to use and accessible via a web interface, making it practical for a broad audience.

Limitations:

- The system struggles with certain car features, especially rare or highly customized vehicles.
- It does not account for regional variations in car prices, which could lead to inaccurate predictions in some cases.
- The model's performance is dependent on the quality and completeness of the input data, which may vary in real-world scenarios.

Future Improvements

Several improvements could be made to enhance the system's accuracy and robustness:

- **Incorporating More Features**: Additional features, such as car service history, accident records, and geographic location, could be added to the model to improve prediction accuracy.
- **Handling Outliers**: The system could include specialized models or techniques for handling outliers, especially for luxury cars or heavily modified vehicles.
- **Location-Based Pricing**: Incorporating regional pricing trends and demand could further improve the model's performance, especially for cars in different geographic locations.
- **Better Data Collection**: More comprehensive data sources should be utilized to ensure that the training dataset is diverse and covers a wider variety of car types and conditions.

The proposed system for predicting used car prices using Artificial Neural Networks and machine learning techniques has proven to be effective in providing accurate price predictions. While it shows promising results, there are areas for improvement, such as handling outliers, incorporating regional price variations, and adding more relevant features. Overall, the system offers a practical and data-driven approach to determining used car prices, which could benefit a wide range of users in the automotive market.

CONCLUSION

This project successfully demonstrated the potential of advanced machine learning techniques in predicting the prices of used cars based on various attributes. By employing an Artificial Neural Network (ANN) model, the system was able to provide accurate price predictions with a high degree of reliability, as evidenced by the strong R^2 value of 0.85 and low Mean Absolute Error (MAE) of approximately $450. This suggests that the model can predict used car prices with a reasonable level of accuracy in most cases. The system was trained using a diverse dataset containing several important features such as make, model, year of manufacture, mileage, and fuel type, which contributed significantly to the predictive power of the model. Through thorough evaluation and testing, the system proved to be valuable in providing quick, data-driven insights into the expected market value of used cars. This is a crucial tool for individuals looking to buy or sell used cars, as well as for dealerships and appraisers who require accurate pricing models for inventory assessment. Despite the promising results, the project identified a few limitations that warrant attention for further improvement. The model's performance was slightly affected by rare car features, and predictions for cars with unusual modifications or from specific geographic regions showed some variation. These challenges highlight the need for better data handling and the inclusion of additional features such as service history and location-based pricing. In conclusion, this project has demonstrated the power of machine learning in addressing a real-world problem, with the potential to significantly improve the accuracy and efficiency of used car price predictions. By enhancing the dataset, refining the model, and incorporating more dynamic factors, the system can evolve into an even more reliable and comprehensive tool for the used car market. The successful implementation of this predictive model opens up further avenues for exploring the application of machine learning in various sectors, showcasing its ability to provide valuable insights and optimize decision-making processes.

REFERENCE:

[1] Tambi, V. K., & Singh, N. Evaluation of Web Services using Various Metrics for Mobile Environments and Multimedia Conferences based on SOAP and REST Principles.

[2] Kumar, T. V. (2024). A Comparison of SQL and NO-SQL Database Management Systems for Unstructured Data.

[3] Kumar, T. V. (2024). A Comprehensive Empirical Study Determining Practitioners' Views on Docker Development Difficulties: Stack Overflow Analysis.

[4] Kumar, T. V. (2024). Developments and Uses of Generative Artificial Intelligence and Present Experimental Data on the Impact on Productivity Applying Artificial Intelligence that is Generative.

[5] Kumar, T. V. (2024). A New Framework and Performance Assessment Method for Distributed Deep Neural NetworkBased Middleware for Cyberattack Detection in the Smart IoT Ecosystem.

[6] Sharma, S., & Dutta, N. (2024). Examining ChatGPT's and Other Models' Potential to Improve the Security Environment using Generative AI for Cybersecurity.

[7] Tambi, V. K., & Singh, N. (2019). Development of a Project Risk Management System based on Industry 4.0 Technology and its Practical Implications. *Development*, *7*(11).

[8] Tambi, V. K., & Singh, N. Blockchain Technology and Cybersecurity Utilisation in New Smart City Applications.

[9] Arora, P., & Bhardwaj, S. Mitigating the Security Issues and Challenges in the Internet of Things (IOT) Framework for Enhanced Security.

[10] Arora, P., & Bhardwaj, S. (2017). A Very Safe and Effective Way to Protect Privacy in Cloud Data Storage Configurations.

[11] Arora, P., & Bhardwaj, S. (2019). The Suitability of Different Cybersecurity Services to Stop Smart Home Attacks.

[12] Arora, P., & Bhardwaj, S. (2020). Research on Cybersecurity Issues and Solutions for Intelligent Transportation Systems.

[13] Arora, P., & Bhardwaj, S. (2021). Methods for Threat and Risk Assessment and Mitigation to Improve Security in the Automotive Sector. *Methods*, *8*(2).

[14] Arora, P., & Bhardwaj, S. Research on Various Security Techniques for Data Protection in Cloud Computing with Cryptography Structures.

[15] Arora, P., & Bhardwaj, S. Examining Cloud Computing Data Confidentiality Techniques to Achieve Higher Security in Cloud Storage.

[16] Arora, P., & Bhardwaj, S. Designs for Secure and Reliable Intrusion Detection Systems using Artificial Intelligence Techniques.

[17] Shreyas, S. K., Katgar, S., Ramaji, M., Goudar, Y., & Srikanteswara, R. (2017). Efficient Food Storage Using Sensors, Android and IoT. *Student BE, Department of CS&E Assistant Professor, Department of CS&E, ramya. srikanteswara@ nmit. ac. in NitteMeenakshi Institute Of Technology, Bengaluru*.

[18] Srikanteswara, R., Reddy, M. C., Himateja, M., & Kumar, K. M. (2022). Object detection and voice guidance for the visually impaired using a smart app. In *Recent Advances in Artificial Intelligence and Data Engineering: Select Proceedings of AIDE 2020* (pp. 133-144). Springer Singapore.

[19] Srikanteswara, R., Hegde, A., & Hegde, P. T. (2023, October). Emergency Vehicle Recognition Via Automated Smart Traffic Manager. In *2023 International Conference on Evolutionary Algorithms and Soft Computing Techniques (EASCT)* (pp. 1-6). IEEE.

[20] Babu, P., Habelalmateen, M. I., Srikanteswara, R., Reddy, R. A., & Purushotham, N. (2024, May). Wafer Surface Semiconductor Defect Classification Using Convolution Neural Network Based Improved Faster R-CNN. In *2024 Second International Conference on Data Science and Information System (ICDSIS)* (pp. 1-4). IEEE.

[21] Srikanteswara, R., & Ramachandra, A. C. (2024). ACM technique for recognition of region of interest using contour and colour features. *Multimedia Tools and Applications*, 1-13.

[22] Srikanteswara, R., Rahul, C. J., Sainath, G., Jaswanth, M. H., & Sharma, V. N. (2022). IoT Based Automatic Medicine Reminder. In *Expert Clouds and Applications: Proceedings of ICOECA 2022* (pp. 157-171). Singapore: Springer Nature Singapore.

[23] Bhat, S. (2015). Design and Function of a Gas Turbine Range Extender for Hybrid Vehicles.

[24] Bhat, S. Automobile Cabin Pre-Conditioning Method Driven by Environmental Conditions with Multi-Satisfaction Goals.

[25] Bhat, S. (2015). Deep Reinforcement Learning for Energy-Saving Thermal Comfort Management in Intelligent Structures.

[26] Bhat, S. (2016). Improving Data Centre Energy Efficiency with End-To-End Cooling Modelling and Optimisation.

[27] Bhat, S. (2020). Enhancing Data Centre Energy Efficiency with Modelling and Optimisation of End-To-End Cooling.

[28] Bhat, S. (2024). Building Thermal Comforts with Various HVAC Systems and Optimum Conditions.

[29] Bhat, S (2023). Discovering the Attractiveness of Hydrogen-Fuelled Gas Turbines in Future Energy Systems.

[30] Bhat, S. (2015). Technology for Chemical Industry Mixing and Processing. *Technology*, *2*(2).

[31] Bhat, S. (2018). The Impact of Data Centre Cooling Technology on Turbo-Mode Efficiency.

[32] Bhat, S. (2019). Data Centre Cooling Technology's Effect on Turbo-Mode Efficiency.

[33] Sharma, S., & Dutta, N. REST Web Service Description Assessment for Hypermedia-Focused Graph-based Service Discovery.

[34] Sharma, S., & Dutta, N. (2017). Development of Attractive Protection through Cyberattack Moderation and Traffic Impact Analysis for Connected Automated Vehicles. *Development*, *4*(2).

[35] Sharma, S., & Dutta, N. (2017). Classification and Feature Extraction in Artificial Intelligence-based Threat Detection using Analysing Methods.

[36] Sharma, S., & Dutta, N. (2018). Development of New Smart City Applications using Blockchain Technology and Cybersecurity Utilisation. *Development*, *7*(11).

[37] Sharma, S., & Dutta, N. (2016). Analysing Anomaly Process Detection using Classification Methods and Negative Selection Algorithms.

[38] Sharma, S., & Dutta, N. (2015). Evaluation of REST Web Service Descriptions for Graph-based Service Discovery with a Hypermedia Focus. *Evaluation*, *2*(5).

[39] Sharma, S., & Dutta, N. (2015). Cybersecurity Vulnerability Management using Novel Artificial Intelligence and Machine Learning Techniques.

[40] Sharma, S., & Dutta, N. (2015). Distributed DNN-based Middleware for Cyberattack Detection in the Smart IOT Ecosystem: A Novel Framework and Performance Evaluation Technique.

[41] Sharma, S., & Dutta, N. (2024). Examining ChatGPT's and Other Models' Potential to Improve the Security Environment using Generative AI for Cybersecurity.

[42] Polamarasetti, A. (2024, November). Research developments, trends and challenges on the rise of machine learning for detection and classification of malware. In 2024 International Conference on Intelligent Computing and Emerging Communication Technologies (ICEC) (pp. 1-5). IEEE.

[43] Polamarasetti, A. (2024, November). Machine learning techniques analysis to Efficient resource provisioning for elastic cloud services. In 2024 International Conference on Intelligent Computing and Emerging Communication Technologies (ICEC) (pp. 1-6). IEEE.

[44] Nagarani, N., et al. "Self-attention based progressive generative adversarial network optimized with momentum search optimization algorithm for classification of brain tumor on MRI image." *Biomedical Signal Processing and Control* 88 (2024): 105597.

[45] Reka, R., R. Karthick, R. Saravana Ram, and Gurkirpal Singh. "Multi head self-attention gated graph convolutional network based multi-attack intrusion detection in MANET." *Computers & Security* 136 (2024): 103526.

[46] Meenalochini, P., R. Karthick, and E. Sakthivel. "An Efficient Control Strategy for an Extended Switched Coupled Inductor Quasi-Z-Source Inverter for 3 Φ Grid Connected System." *Journal of Circuits, Systems and Computers* 32.11 (2023): 2450011.

[47] Karthick, R., et al. "An optimal partitioning and floor planning for VLSI circuit design based on a hybrid bio-inspired whale optimization and adaptive bird swarm optimization (WO-ABSO) algorithm." *Journal of Circuits, Systems and Computers* 32.08 (2023): 2350273.

[48] Jasper Gnana Chandran, J., et al. "Dual-channel capsule generative adversarial network optimized with golden eagle optimization for pediatric bone age assessment from hand X-ray image." *International Journal of Pattern Recognition and Artificial Intelligence* 37.02 (2023): 2354001.

[49] Rajagopal RK, Karthick R, Meenalochini P, Kalaichelvi T. Deep Convolutional Spiking Neural Network optimized with Arithmetic optimization algorithm for lung disease

detection using chest X-ray images. Biomedical Signal Processing and Control. 2023 Jan 1;79:104197.

[50] Karthick, R., and P. Meenalochini. "Implementation of data cache block (DCB) in shared processor using field-programmable gate array (FPGA)." *Journal of the National Science Foundation of Sri Lanka* 48.4 (2020).

[51] Karthick, R., A. Senthilselvi, P. Meenalochini, and S. Senthil Pandi. "Design and analysis of linear phase finite impulse response filter using water strider optimization algorithm in FPGA." *Circuits, Systems, and Signal Processing* 41, no. 9 (2022): 5254-5282.

[52] Karthick, R., and M. Sundararajan. "SPIDER-based out-of-order execution scheme for Ht-MPSOC." *International Journal of Advanced Intelligence paradigms* 19.1 (2021): 28-41.

[53] Karthick, R., Dawood, M.S. & Meenalochini, P. Analysis of vital signs using remote photoplethysmography (RPPG). *J Ambient Intell Human Comput* **14**, 16729–16736 (2023). https://doi.org/10.1007/s12652-023-04683-w

[54] Polamarasetti, A. (2024, November). Role of Artificial Intelligence and Machine Learning to Enhancing Cloud Security. In 2024 International Conference on Intelligent Computing and Emerging Communication Technologies (ICEC) (pp. 1-6). IEEE.

[55] Vadisetty, R., & Polamarasetti, A. (2024, November). Quantum Computing For Cryptographic Security With Artificial Intelligence. In 2024 12th International Conference on Control, Mechatronics and Automation (ICCMA) (pp. 252-260). IEEE.

[56] Vadisetty, R., & Polamarasetti, A. (2024, November). Generative AI for Cyber Threat Simulation and Defense. In 2024 12th International Conference on Control, Mechatronics and Automation (ICCMA) (pp. 272-279). IEEE.

[57] Mishra, R., Saurabh, S., Dwivedi, S., & Singh, V. (2025). Influential Social Media Marketing by Integrating the Strategic Implementation. In Data Analytics and Influencer Marketing for Cultivating Brand Evangelism and Affinity (pp. 411-430). IGI Global Scientific Publishing.

[58] Sharma, A., Aggarwal, M., Singh, V., Seth, K., & Thakur, S. (2001). Post-Covid 19 essential skills: A study of tourism and hospitality graduates. Operational transformations in tourism and hospitality, 157.