LITERATURE SURVEY ON CAR RETAIL PRICE PREDICTION

PUBLICATION / YEAR	TOPIC	OVERVIEW	POSITIVE ASPECTS	LIMITATIONS
AEA / Feb 2013	Evidence from New and Used Car Purchases†	Hausman (1979) was the first to investigate whether consumers are myopic when purchasing durable goods that vary in energy costs. More generally. These estimates become the building blocks for our next step. We find little evidence that consumers "undervalue" future gasoline costs when purchasing cars.	In this paper we have estimated the effect of gasoline prices on the short-run equilibrium prices, market shares, and sales of new and used cars of different fuel economies. We also estimated the effect of a $1 increase in gasoline prices on unit sales of new cars and found that sales in the highest fuel economy quartile increased by 10–12 percent, while sales in the lowest fuel economy quartile fell by 27–28 percent	We have used these estimates to address a question that is important for understanding the ability of a policy intervention such as a gasoline tax or a carbon tax to influence what cars people buy. This could be thought of as a necessary condition for effective policy: the more car buyers discount future fuel costs, the less effective a gasoline tax or carbon tax will be in influencing vehicle choice.
LCR / 27 Oct 2021	Object Detection and Used Car Price Predicting Analysis System (UCPAS) Using Machine Learning Technique	The highly interesting research area that noticed in the last few years is object detection and find out the prediction based on the features that can be benefited to consumers and the industry. The outcome of this experiment shows that clustering with linear regression and Random Forest model yield the best accuracy outcomes	The huge number of challenges for the Used car price predicting analysis system (UCPAS) such as the large number of parameters that are considered during the prediction process, if the system selected the wrong parameter, it will drastically affect the outcomes. The pre-processing is required to increases the performance of UCPAS.The R2 score of Regression analysis was good for predictions and close to the original selling prices in the market.	In this paper, Python is used for pre-processing (major step of UCPAS) to normalize and data cleaning (remove ambiguity, duplicates and noise). The pre-processing is required to increases the performance of UCPAS. the disadvantage of the proposed UCPA System is that it required a large number of features considered and the huge number of computational tasks.

NSFC / 24 July 2020	Deep end-to-end learning for price prediction of second-hand items	Recent years have witnessed the rapid development of online shopping and ecommerce websites, e.g., eBay and OLX. Online shopping markets offer millions of products for sale each day. State-of-the-art methods can predict the price of only one item category. This proposed method utilizes a deep neural network involving long short-term memory (LSTM) and convolutional neural network architectures for price prediction. Using a dataset crawled from a website for second-hand items, the proposed method of combining the predicted secondhand item quality score with the forecasted minimum and maximum price outperforms the other models in all of the used accuracy metrics with a significant performance gap.	This work has two main contributions. The first contribution is to predict the price of secondhand items based on textual and visual features for different product types, as state-of-the-art ecommerce price prediction methods do not focus on visual features. The proposed model highlights the feasibility of combining images and textual data to make a prediction. Additionally, the proposed method allows a single model to be applicable to a dataset of different product types, in contrast to other price prediction models that use different prediction models to handle datasets of different categories of products.The MAE scores of the proposed price prediction model and the SVM baseline model are 0.07 and 0.09, respectively.	We utilized a hybrid CNN-LSTM model for the task of price prediction which achieved a better performance in comparison with the baseline model. The MAE scores of the proposed price prediction model and the SVM baseline model are 0.07 and 0.09, respectively. The second contribution is to improve the predicted price of the first contribution. Thus, we proposed forecasting the minimum and maximum prices of the second-hand item, i.e., the price range of the product type this item belongs to.We utilized a hybrid CNN-LSTM model for the task of price prediction which achieved a better performance in comparison with the baseline model.

PSO-GRA-BP Neural Network / 22 July 2022	Research on the Prediction Model of the Used Car Price	As the mobile Internet improves by leaps and bounds, the model of traditional offline used car trading has gradually lost the ability to live up to the needs of consumers, and online used car trading platforms have emerged as the times require. In order to standardize the evaluation standards of used car prices and improve the accuracy of used car price forecasts, the linear correlation between vehicle parameters, vehicle conditions, and transaction factors and used car price was comprehensively investigated, grey relational analysis was applied to filter the feature variables of factors affecting used car price, the traditional BP neural network was also optimized by combining the particle swarm optimization algorithm, and a used car price prediction method based on PSO-GRA-BPNN was proposed.	In recent years, online used car trading platforms have developed rapidly, but they still face many problems. In practice, institutions and individuals differ in how they screen the characteristic variables of used car prices and predict used car prices. Using web crawler technology to obtain used car transaction data, three prediction models of BPNN, GRABPNN, and PSO-GRA-BPNN were constructed to conduct comparative verification and result analysis. In a rough comparison, the BPNN model has lower accuracy, with an error range of about 19.979%, and it is unstable. Although the PSO-GRA-BPNN used car price prediction model has high prediction accuracy, it has lost time. It is mainly analyzed from two aspects.	In terms of theory, traditional used car price evaluation methods rely too much on the subjective judgment of evaluators, which can no longer meet the needs of online transactions in the used car market. Therefore, it is necessary to establish an efficient, reasonable, fair, and accurate used car price evaluation system. In a rough comparison, the BPNN model has lower accuracy, with an error range of about 19.979%, and it is unstable. In the case of feature variable screening, the prediction accuracy of the GRA-BPNN model is higher than that of the BPNN, and the error range is about 13.986%. In addition, in view of the fact that the prediction accuracy of high-end used cars is lower than that of low-end used cars, it is suggested that when pricing high-end used cars, you need to check other configuration information in order to make a more reasonable judgment.

JCSE / 31 May 2019	Car Price Prediction Using Machine Learning	Because of new computing technologies, machine learning today is not like machine learning of the past. It was born from pattern recognition and the theory that computers can learn without being programmed to perform specific tasks; researchers interested in artificial intelligence wanted to see if computers could learn from data. It's a science that's not new – but one that has gained fresh momentum. While there is an end number of applications of machine learning in real life one of the most prominent application is the prediction problem. There are various topics on which the prediction can be applied. One such application is what this project is focused upon. This ability to capture data, analyze it and use it to personalize a shopping experience (or implement a marketing campaign) is the future of retail.	In this chapter, we discuss the results and observation we did while implementing this module. We successfully implemented the machine learning algorithmic paradigms using prominent algorithms from libraries in python. We first perform pre-processing and data cleaning on our dataset. Positive correlation basically relates to the concept of direct proportion whereas Negative correlation relates to the concept of inverse proportion. K Nearest Neighbour (KNN) and Classification and Regression Trees (CART) are compared on two different models of vehicles.	We found that 15% of the tuples had null values and we pruned those tuples. The results showed that there is a positive correlation between price and kilometers traveled, year of registration and kilometers traveled and a negative correlation between price and year of registration. The year of registration was slightly more dominant. As we know technologies are improving day by day and there is also advancement in car technology also, so our next upgrade will include hybrid cars, electric cars, and Driverless cars.