Industrial Engineering Journal
ISSN: 0970-2555
Volume : 52, Issue 9, September : 2023

USING MACHINE LEARNING AND NEURAL NETWORKS TO PREDICT THE VALUE OF USED VEHICLES

[#1]CHEKURTHI LATHA,
[#2]Dr. P.VENKATESHWARLU, ***Associate Professor,***
[#3]Dr. V.BAPUJI, ***Associate Professor& HOD,***
Department of Master of Computer Applications,
VAAGESWARI COLLEGE OF ENGINEERING, KARIMNAGAR, TELANGANA

ABSTRACT: Choosing the right price when purchasing or selling a car can be difficult. For such calculations, artificial neural networks, a subtype of artificial intelligence, are frequently used. We built two distinct artificial neural networks for predicting automobile pricing and evaluated them using data from a car sales website in this study. A C# software was written to read the data on the website, and it made use of the MSSQL Server database management system and the HTMLAgilityPack package. To digitize the data, a technique in the T- SQL language was built. The study, which assessed data from about 1000 automobiles, had a 91.38% success rate. More data is required for better results.

Keywords: Artificial Intelligence, Artificial NeuralNetworks, Car Price Prediction, HTMLAgilityPack.

1. INTRODUCTION

To solve real-world problems, artificial neural networks (ANNs) attempt to mimic human cognitive processes. Classification, prediction, linking, sorting, and understanding data are only few of the common uses for artificial neural networks (Ayar, 2015). With the advancement of technology, ANNs have found increasing application. There is a wide range of applications for artificial neural networks. The military employs them for a variety of purposes, including image processing, simulation, and social event forecasting. These networks have several applications in healthcare, including disease diagnosis and treatment, disease prevention, pharmaceutical study, and patient isolation. Bourquin Schmidt and Van Hooghees have devoted a great deal of time to studying the potential of simulation technologies in a variety of fields, including the space sciences, banking and finance, the prediction of stock prices, the forecasting of economic crises, the development of management games, and the forecasting of currency exchange rates.

The economic value of the automotive industry is believed to be in the billions of dollars due to the large number of people who have bought and sold cars at some point in their lives. This highlights the significance of striking a deal that benefits both the buyer and the seller equally. Estimates are typically arrived at by comparing the asking price to the prices of recently sold cars that are quite comparable to the one being considered. Using a collection of 20 variables, this research aimed to determine car prices using artificial neural networks.

The prices of comparable vehicles are often considered by buyers and sellers. Sahibinden.com is widely regarded in Turkey as one of the greatest sites for this niche. It is standard practice for purchasers to propose a greater asking price than they anticipate receiving when reselling an item. If someone else with a comparable vehicle sees this data, they may increase the asking price

Industrial Engineering Journal
ISSN: 0970-2555
Volume : 52, Issue 9, September : 2023

of their own vehicle in the mistaken belief that it is of higher quality. Daştan (2016) claims that this practice significantly raises the cost of automobiles. Car production dropped during the COVID-19 epidemic, driving up the price of this product and creating a demand all on its own. Several sources (Koc & Bilgic, 2020; Ekonomist, 2020; Aymelek, 2020) claim that in Turkey, the price of a car is consistently higher than inflation.

An AI-powered tool for calculating costs is likely to be developed in the not-too-distant future. To maintain a level playing field for buyers and sellers, this application will analyze how the aforementioned procedures influence current prices. By contrasting the prices offered by automobile dealerships with those projected by software and those listed by vendors on auto sales websites (Daştan, 2016; Agarwal & Ratchford, 1980), the upcoming study will investigate the impact of linked websites on price hikes.

2. RELATED WORK

According to research done in 2005 by Talu and Tatar, an artificial neural network (ANN) is a model of a computer system that is inspired by the structure and function of the human nervous system. Figure 1 is an example of a simplified artificial neural network (ANN).

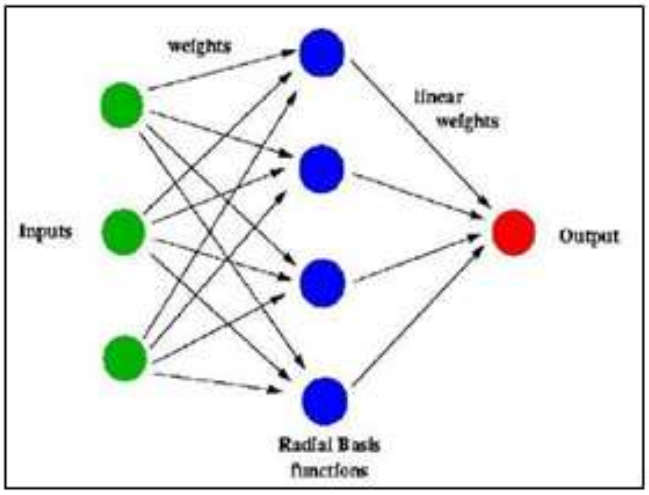

Fig. 1 The basic block diagram of an artificial neural net (Laxmi & Rohil, 2014).

The two most crucial components of an Artificial Neural Network (ANN) are the learning method and the activation function. The learning algorithm employs real-world data to determine optimal weights, allowing for optimal performance. The activation function plays a crucial role by transforming the incoming data into a predetermined range. Multiple layers of neurons make up artificial neural networks (ANNs). (Han and Kopacck, 1995) Bourquin, Schmidli, Van Hoogevest, and Leuenberger (1998) and Sieniutycz (2019) cite the aforementioned works.

A neuron is a specialized nerve cell, and neural networks are made up of billions of them. Each neuron receives and sends a unique set of numbers. Connections between a cell and its external environment or other cells are established by means of weights. Finding the final sum is as simple as using addition. The final output is calculated by applying the activation function to the resulting number. This factor is responsible for the output of the cell's total number (Han and Kopacck, 1995). The neuron in Figure 2 is depicted in block form. Displays the interdependence of neuronal components.

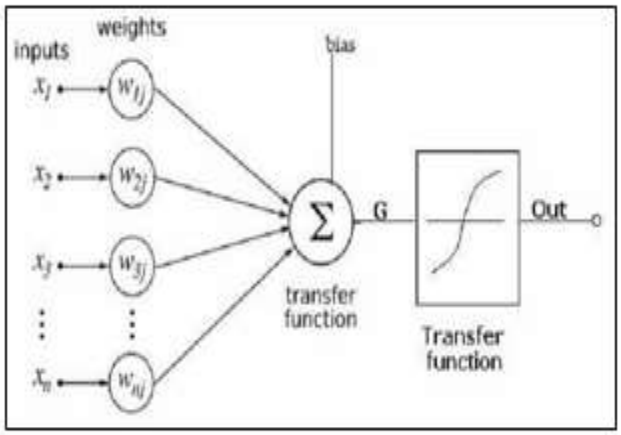

Fig. 2 The basic block diagram of a neuron (Thati,Biswas, & Chowdhury, 2015).

Artificial Neural Networks (ANNs) have many potential applications; some of them include expert systems, optimization, classification, pattern recognition, and control (Thati, Biswas, and Chowdhury, 2015; Deperliolu and Kose, 2011; Elmas, 2003). Elmas (2003) states that the typical construction of an ANN consists of three

layers of neurons, but that this architecture might shift depending on the problem domain.

Input Layer

An artificial neural network (ANN)'s input layer is responsible for gathering data from the external environment. The inputs are transmitted directly to the hidden layer, which has the same amount of cells, without going through any preprocessing layers.

Hidden Layers

Through hidden levels, data from the input layer is passed on to the subsequent layer. In neural networks, the number of cells in each hidden layer and the order in which they are hidden can vary widely amongst networks. The provided numbers to the subterranean layer are independent of the input and output numbers.

Output Layer

Data transferred from the secret layer is received by the output layer, where it is transformed, and then sent on to the rest of the world. It's possible that the output layer contains more than one cell. Each layer's output cells are tightly interconnected and each layer's output cells have a single output.

3. SYSTEM DESIGN

To properly train and evaluate a neural network simulator, a dataset is required (Elmas, 2003). According to Google Trends (n.d.), sahibinden.com is the most visited car-selling website in Turkey, and this is where the advertisements used in this study were sourced. The database management system and data visualization tool utilized were developed using C# and Microsoft's MSSQL Server. The data retrieval from the site was accomplished with the help of HTMLAgilityPack. As a result, information from 1067 vehicles was included. The data was initially converted to digital format. The data then required normalization. This was done to ensure that the ANN was functioning properly and that the weights within the ANN were uniform. In 2017, Bala and Kumar were tasked with correcting inaccurate and incomplete data in order to generate usable statistics.

A multi-layered feedback artificial neural network (ANN) was constructed in MATLAB with 20 inputs (representing different parameters) and a single output (representing price) to facilitate prediction. Brand, series, model, year, fuel type, gearbox type, mileage, body style, motor power, engine displacement, traction system, color, warranty status, damage condition, license plate information, origin, trade data, and the state of the vehicle were all input categories. An infinite number of possible values could be included in this set. Sedans, station wagons, five-door hatchbacks, coupes, three-door hatchbacks, minivans, cabriolets, and crossovers are all included in the body type metric. The range of values for the parameters is displayed in Table 1. The cost reflected the value of the data.

Table 1 Diversity of input values

Inputs	Diversity of Value
Brand	32
Series	153
Model	477
Year	Didn't replaced
Fuel	3
Gear	3
Km	Didn't replaced
Body type	8
Motor power	107
Engine displacement	95
Traction	4
Color	18
Warranty	2
Damage condition	4
Plate	2
From	2
Trade	2
State	2

There were two ANNs built, each with a unique combination of hidden layers and neurons. Some of it was utilized for instruction, and some of it

Industrial Engineering Journal
ISSN: 0970-2555
Volume : 52, Issue 9, September : 2023

was put to the test. As can be seen in Figure 3, the original ANN consisted of two hidden layers, the first of which included 30 neurons and the second of which contained 25. To train the model, 70% of the data was used, while the remaining 30% was used for testing.

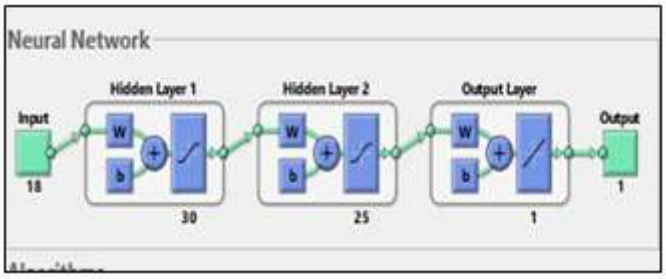

Fig. 3 ANN (Two Hidden Layers)

The neutron counts for the second ANN are shown in Figure 4 to be 30, 15, and 25. About 60% of the data was put to use in the training process, while the remaining 40% was put to use in the testing phase.

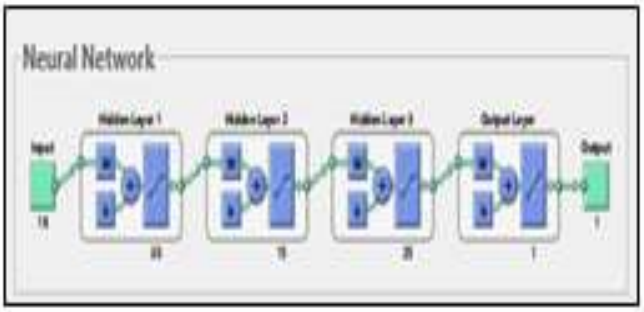

Fig. 4 ANN (Three Hidden Layers).

4. RESULTS

An ANN with two hidden layers containing 30 and 25 neurons is depicted in Figure 5's performance curve.

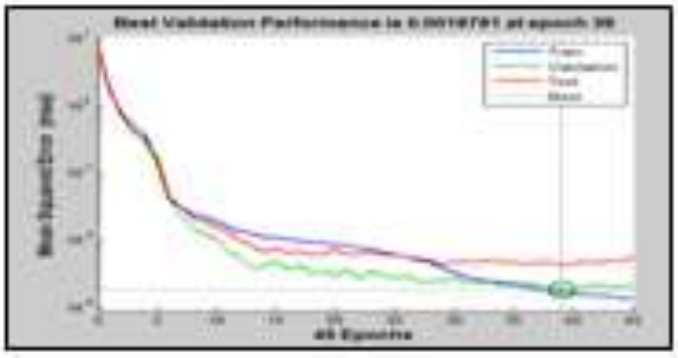

Fig. 5 Success Chart of ANN (Two Hidden Layers).

As can be seen in Figure 5, the program was run 45 times, with a minimum squared error of 0.0018781 after the first 39 runs. The program was terminated since the problem was becoming worse with each repetition. Figure 6 demonstrates the ANN's impressive performance.

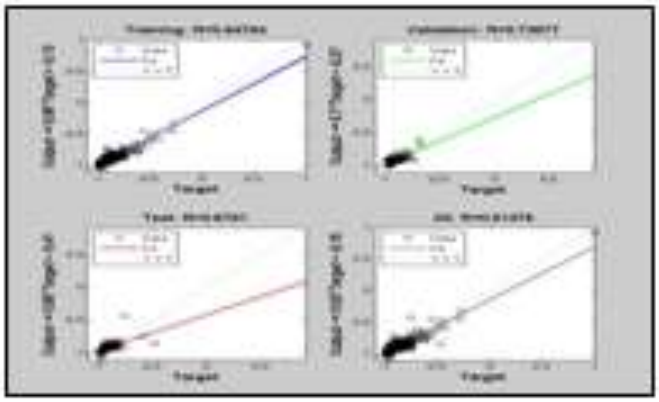

Fig. 6 The performance of the first ANN.

In Figure 7, we see a simplified block diagram of an artificial neural network (ANN). Each of the ANN's three secret layers is equipped with 30, 15, and 25 neurons.

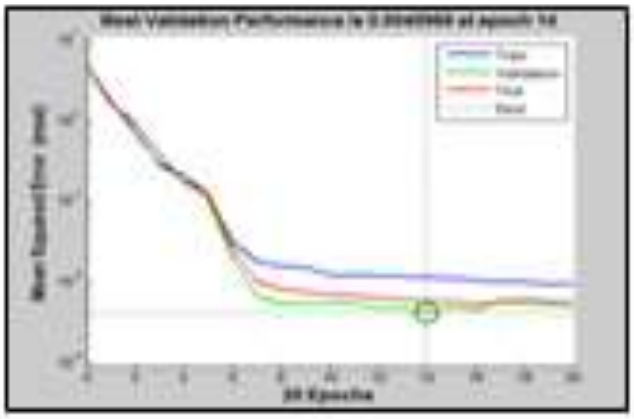

Fig. 7 Success Chart of ANN (three Hidden Layers).

Figure 7 displays the results of 20 program runs, with a squared error estimate of 0.0040969 after 14 runs. Error rates were rising with each cycle, therefore the software was shut down. Figure 8 demonstrates the ANN's efficacy.

Industrial Engineering Journal
ISSN: 0970-2555
Volume : 52, Issue 9, September : 2023

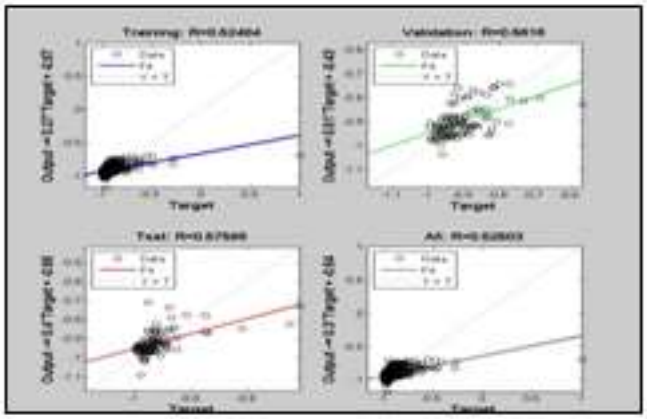

Fig. 8 The performance of the second ANN.

Training with less data results in more levels and more neurons but lower accuracy. The study found that participants' estimates of car prices were accurate 91.38 percent of the time. A significantly larger data collection is required for more precise results. Due to security concerns, the website only allowed a limited amount of requests, thus further details could not be obtained. However, the predictions are in line with common assumptions. The next analysis will compare the prices that buyers and sellers actually enter onto the website with the prices that the software forecasts. This will demonstrate the effect of the linked site on cost rises.

REFERENCES

1. AGARWAL, M. K., & RATCHFORD, B. T. (1980). Estimating Demand Function for Product. Journal of Consumer Research, 7(3), 249- 262.
2. AĞYAR, Z. (2015). Use of Artificial Neural Networks (in Turkish). Engıneer and Machine, 56(662), 22-23.
3. BALA, R., & KUMAR, D. (2017). Classification Using ANN: A Review. International Journal of Computational Intelligence Research, 13(7), 1811-1820.
4. BOHR, H. (2020). Drug discovery and molecular modeling using artificial intelligence. In A. Bohr, & K. Memarzadeh, Artificial Intelligence in Healthcare (pp. 61-83). Academic Press.
5. DAŞTAN, H. (2016). Determination of The Factors That Effect Second-Hand Automobile Prices in Turkey by Using Hedonic Pricing Model (in Turkish). Gazi Üniversitesi İktisadi ve İdari Bilimler Fakültesi Dergisi, 1(18), 303- 327.
6. HAN, M.-W., & KOPACCK, P. (1995). Optimization Using Artificial Neural Networks. IFAC Proceedings Volumes, 28(24), 357-360.
7. KARCI, A., & DEMIR, M. (2009). Estimation of protein structures by classification of angles between a carbons of amino acids based on artificial neural Networks. Expert Systems with Applications, 36(1), 5541-5548.
8. KOC, C., & BILGIC, T. (2020, February 15). New- car shortages lead Turkish run on used vehicles. Retrieved from Autonews.com: https://europe.autonews.com/sales-market/new-car-shortages-lead-turkish-run-used-vehicles
9. LAXMI, V., & ROHIL, H. (2014). License Plate Recognition System using Back Propagation Neural Network. International Journal of Computer Applications, 99(8), 29-37.
10. LI, B., LEE, Y., YAO, W., LU, Y., & FAN, X. (2020). Development and application of ANN model for property prediction of supercritical kerosene. Computers and Fluids(209), 1-18.
11. OZSAHIN, İ., & OZSAHIN, D. U. (2020). Neural network applications in medicine. In W. Zgallai, Biomedical Signal Processing and Artificial Intelligence in Healthcare (pp. 183-206). Academic Press.
12. SIENIUTYCZ, S. (2019). Complex systems of neural networks. In S. Sieniutycz, Complexity and Complex Thermo-Economic Systems (pp. 51- 84). Elsevier.
13. TALU, M. F., & TATAR, Y. (2005). An Edge detection Method for Using Model based Artificial Neural Network (In Turkish).

Journal of Turkey Informatics and Computer Science Foundation, 1(1).

14. THATI, A., BISWAS, A., & CHOWDHURY, S. R. (2015). Breath acetone-based non-invasive detection of blood glucose levels. International Journal on Smart Sensing and Intelligent Systems, 2(8), 1244-1260.