Fuel Efficiency Prediction Using Deep Learning

NIPUN SINGH
Dept. of Computer Science and Engineering
Lovely Professional University
Jalandhar, INDIA
nipun97531@gmail.com

Abstract

Fuel efficiency plays a pivotal role in modern transportation systems, holding significant implications for both environmental sustainability and economic viability. Precisely estimating a vehicle's fuel efficiency is crucial for manufacturers, consumers, and policymakers alike. Traditional methods for gauging fuel efficiency often rely on standardized testing procedures, which may fall short of accurately representing real-world driving conditions. This research delves into the application of deep learning techniques for enhanced fuel efficiency prediction. By utilizing a broader spectrum of variables and capturing the intricate interplay of factors impacting fuel consumption.

This study harnesses Artificial Neural Networks (ANNs). These ANNs mimic human brain processes through mathematical computations, facilitating various advancements in fields such as artificial intelligence, encompassing image and voice recognition, robotics, and more. ANNs emulate the structure of the human brain, featuring neurons interconnected via weighted links in a complex and nonlinear fashion.

In summary, this research contributes significantly to advancing fuel efficiency prediction by demonstrating the efficacy of deep learning methods in comprehending the multifaceted nature of vehicle fuel consumption. These findings underscore the vital role of artificial intelligence in achieving more accurate predictions, thereby fostering energy conservation, environmental sustainability, and cost savings within the transportation sector.

Keywords: ANN, Transportation System, Artificial Intelligence

1. Introduction

.

Deep learning is indeed a powerful subset of machine learning that excels in handling unstructured data, such as text and images. Unlike traditional machine learning algorithms, deep learning models have the ability to automatically learn and extract relevant features from the data, reducing the need for extensive manual feature engineering.

In the context of image classification, as you mentioned, deep learning systems are particularly effective. They can analyze and understand intricate patterns and features within images without human intervention. For example, when classifying images of pets into categories like "cat," "dog," or "hamster," deep learning models can automatically learn to recognize distinguishing features like fur patterns, shapes, and other visual cues that differentiate these animals.

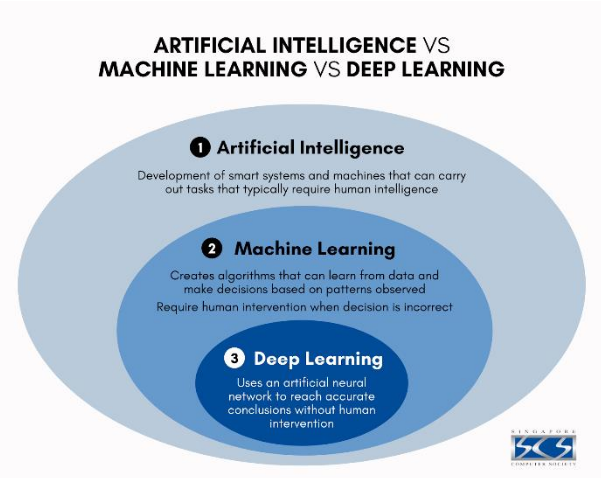


Fig 1: Deep Learning

This ability to handle unstructured data and automate feature extraction makes deep learning well-suited for a wide range of applications beyond image classification. It's used in natural language processing (NLP) for text analysis, speech recognition, recommendation systems, and even in more complex tasks like autonomous driving and medical image analysis.

Deep learning models, such as convolutional neural networks (CNNs) for images and recurrent neural networks (RNNs) for sequential data like text, have achieved remarkable success in various domains. They've significantly advanced the field of artificial intelligence and continue to be a driving force in the development of cutting-edge applications. However, it's essential to consider data quality, model complexity, and computational resources when implementing deep learning solutions, as they can be computationally intensive and require substantial amounts of labeled data for training.

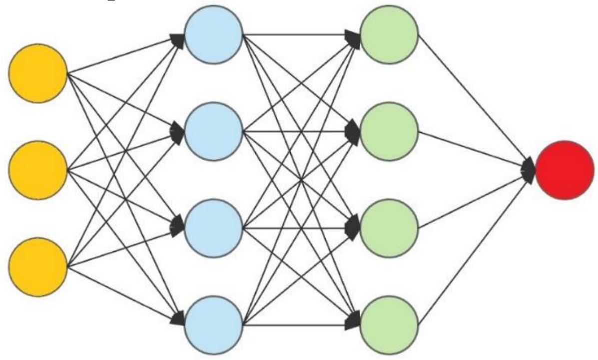


Fig 2: Neural Network layers

2. Importance of Fuel Prediction

Fuel efficiency is of paramount importance in today's world for several compelling reasons. Firstly, it plays a vital role in environmental sustainability. Improved fuel efficiency reduces the amount of greenhouse gas emissions, which are major contributors to climate change. As we strive to combat the adverse effects of global warming and reduce our carbon footprint, more fuel-efficient vehicles and transportation systems are essential.

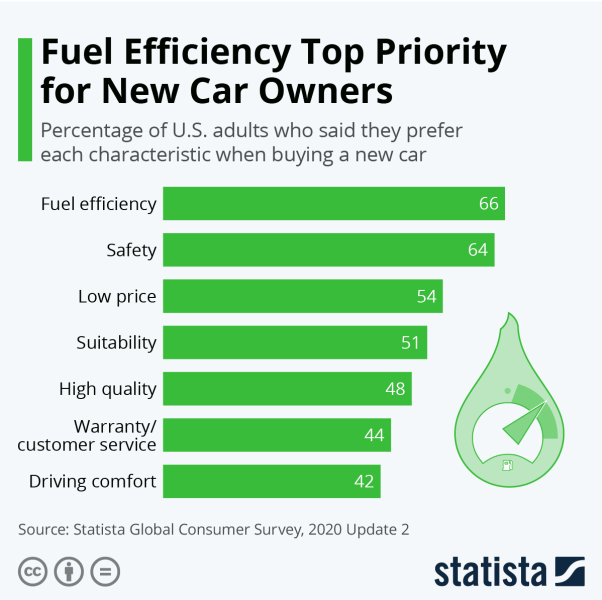


Fig 3: Fuel Efficiency and its importance

Secondly, fuel efficiency has a direct impact on our wallets. Higher fuel efficiency means lower fuel consumption and, consequently, reduced expenses for individuals, families, and businesses. It eases the financial burden of fuel costs, especially during times of fluctuating oil prices.

Additionally, enhanced fuel efficiency contributes to energy security by reducing our dependence on foreign oil imports. This is not only beneficial for national economies but also enhances energy resilience and minimizes the vulnerability of supply disruptions.

From a technological standpoint, advancements in fuel efficiency drive innovation and the development of cleaner and more sustainable transportation solutions. It fosters research and development, creating opportunities for breakthroughs in alternative energy sources and propulsion systems.

In summary, fuel efficiency is a multifaceted concept with far-reaching implications for the environment, economy, and energy security. Embracing and prioritizing fuel efficiency is a proactive step toward a more sustainable and economically sound future.

At its most basic, fuel efficiency is defined as a measure of how much a car will convert energy in fuel into kinetic energy to travel. In other words, fuel efficiency shows how far your car can travel with a certain amount of fuel. In America, the concept is described as “miles per gallon” (mpg). Vehicles with better fuel efficiency tend to consume less fuel to carry out the same task. Therefore, reducing wasted fuel. Choosing a fuel-efficient vehicle can bring a wide range of advantages: saving fuel costs, reducing carbon footprint, cutting our dependence on oil, etc. Let’s take a quick look at why fuel efficiency is a crucial element you need to take into account, as well as the benefits it can offer you as a driver and a responsible citizen alike.

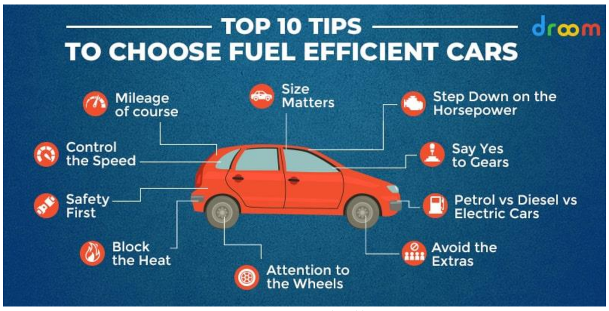


Fig 4: Properties of Efficient cars

Save Money on Gas

Oil prices are one of the major issues facing drivers today. Given the rising prices of fuel, it’s probably time for you to consider saving money on gas. The gas mileage, the amount of gas your car consumes per mile, plays an important role in how much money you save on gas each year. According to some economists, you might save 4,500 USD for 5 years by driving a vehicle that gets 30 mpg rather than 20 mpg. Thus, if you start trying to drive a more fuel-efficient car and use less fuel, not only will you save a lot of money, but you can also spend that saved money on something more meaningful to you instead.

Reduce Carbon Footprint

While you can try plenty of little things in your daily lives to reduce your carbon footprint, driving a car with better fuel efficiency is undoubtedly the best way to fight climate change. A recent study found that driving a more fuel-efficient vehicle is by far the most realistic and effective action to achieve the largest cuts in emissions. The research says a car that gets 30 mpg would reduce total emissions by 5% than a vehicle that gets 20 mpg. So, if you want to contribute to a decrease in greenhouse gas emissions, buying a car that boasts better gas mileage may be the best option.

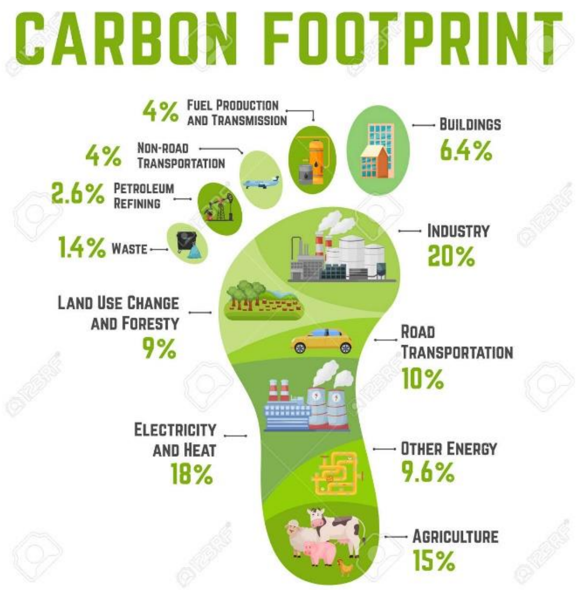


Fig 5: Carbon Footprint

Lower Dependence on Oil

According to statistics, over 70% of total U.S. on-road vehicles consume oil, and they account for nearly a fourth of the country's emissions, contributing to climate change. Moreover, the country paid about 120 million USD in 2014 for overseas oil, which was mainly imported from the Middle East. By owning a fuel-efficient vehicle, you can reduce dependence on oil, as well as save money for both yourself and your country.

Fuel, primarily composed of carbon, is a substance that, when burned, releases a significant amount of thermal energy. Fuels are integral to various domestic and industrial applications, serving as a vital source of energy.

Energy, as a fundamental principle, cannot be created or destroyed but can undergo transformation from one form to another. For instance, when gasoline is combusted in a car engine, its chemical energy is converted into thermal energy, which, in turn, transforms into mechanical energy that powers the vehicle, ultimately manifesting as kinetic energy. Various types of fuels, such as coal, petrol, diesel, and wood, exist, each with specific use cases. Notably, not every fuel can be used interchangeably; their applications depend on their efficiency in a given context.

Fuel efficiency, therefore, is the capacity of a vehicle or any device to effectively harness energy from the fuel it consumes. It measures how well a fuel can convert one form of energy into another. This measurement is typically expressed through the fuel's calorific value.

Calorific value refers to the quantity of energy liberated when 1 kilogram of fuel undergoes complete combustion. It is typically quantified in kilojoules per kilogram (kJ/kg). Fuel efficiency is closely tied to this calorific value. When comparing two fuels of equal quantity, the one with a higher calorific value is considered more efficient, as it can extract more energy from the same amount of fuel.

In summary, fuel serves as a valuable energy source with the ability to transform energy from one state to another. Fuel efficiency, as determined by calorific value, plays a crucial role in evaluating the effectiveness of different fuels in various applications.

3. Construction of ANN model

3.1 Importing important Libraries

Machine learning, while complex, has become more accessible through the use of machine learning frameworks like Google's TensorFlow. TensorFlow, introduced by the Google Brain team in 2015, is an open-source library designed for numerical computation and large-scale machine learning. It simplifies the process of data acquisition, model training, making predictions, and refining results.

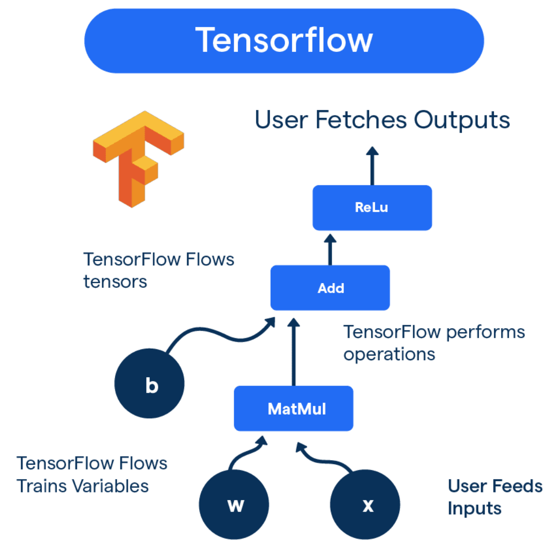


Fig 6: Tensorflow

TensorFlow encompasses a wide range of machine learning and deep learning models and algorithms, including neural networks, and offers a user-friendly programming interface using Python or JavaScript. Behind the scenes, it leverages high-performance C++ for efficient execution.

Competing with frameworks like PyTorch and Apache MXNet, TensorFlow is versatile. It can be used for tasks such as training deep neural networks for tasks like image recognition, natural language processing, and simulations based on partial differential equations. Notably, TensorFlow supports deploying models for large-scale production predictions, ensuring a seamless transition from training to real-world use.

3.2 Network Structure

An Artificial Neural Network (ANN) is composed of artificial neurons organized into three main layers: the input, hidden, and output layers. The input layer is responsible for transferring data from the external world to the hidden layer. Unlike the other layers, the input layer doesn't process the data; it merely serves as a conduit for information. In the hidden layer, outputs are generated by using the data from the input layer, along with bias, summation, and activation functions. ANNs can have more than one hidden layer, with each hidden layer passing its outputs to the subsequent layer.

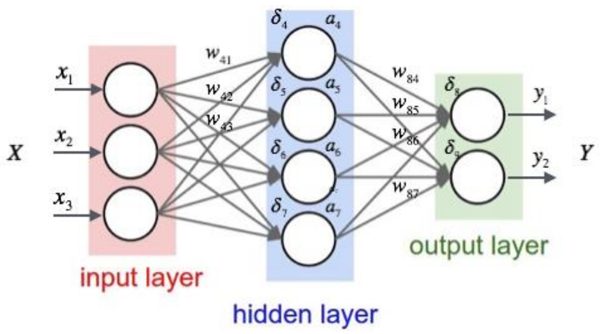


Fig 7: Different layers in Neural Network (notice hidden layeyer)

The output layer produces the network's final output by processing data from the hidden layer and sending it to the external world. The summation function calculates the net input received by a cell, and different functions can be used for this purpose, with the weighted sum being the most common. The inputs to the network's input cells, represented as (N, T, m, Tin, Tcw), represent knowledge from other cells or the external world and are determined by the examples the network aims to learn. The weights (w1, w2, ...wn) determine the impact of the input set or another processing element in the previous layer on the processing element.

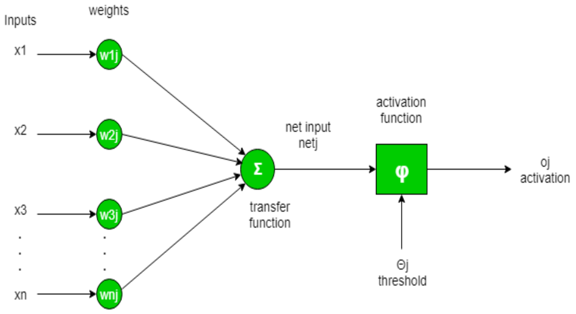
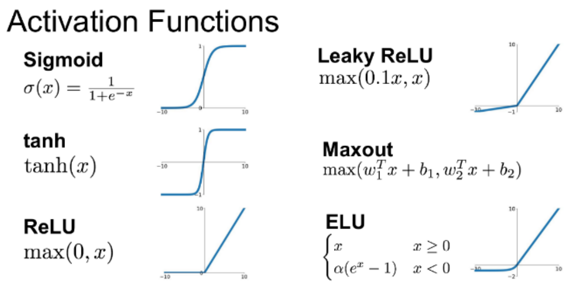


Fig 8: Activation Function

Each input value is multiplied by the weight connecting it to the processing element and then combined using the summation function to find the net input of the network, as shown in Equation (1).

The activation function establishes a non-linear relationship between the input and output layers and determines a cell's output by processing the net input. The choice of an appropriate activation function significantly influences the network's performance. In this study, the logistic sigmoid transfer function is employed as an activation function in the multilayer perception model. This function is favored because it is differentiable, continuous, and non-linear. It produces output values between 0 and 1 for each net input value, as indicated in the logistic sigmoid function formula provided.



3.3 Learning Algorithm

To determine the weights in an artificial neural network (ANN), there are many different learning algorithms. One of the most popular algorithms is back propagation. This algorithm updates the weights of the ANN based on the difference between the desired output and the actual output of the network. The learning parameter used in this algorithm is very important for achieving the best results. The learning parameter can be either constant or dynamically updated.

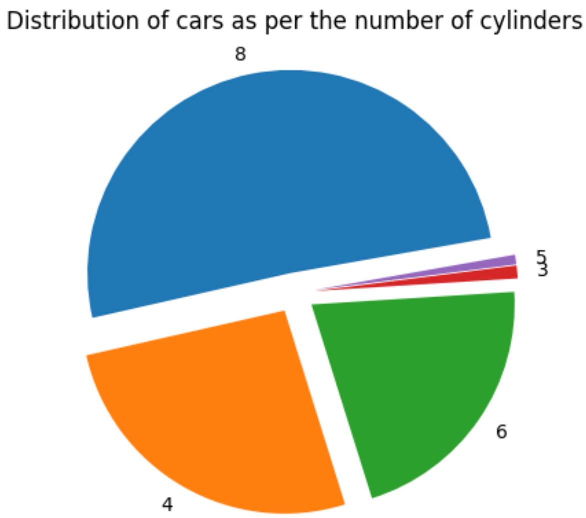


Fig 9: Distribution of cylinders in car

Previous studies have used a variety of training functions, such as Bayesian regularization, gradient descent with adaptive learning rule, gradient descent with momentum and adaptive learning rule, scaled conjugate gradient, and Levenberge-Marquardt. The goal of these training functions is to find the weights that minimize the error between the desired output and the actual output of the network.
In order to find the best learning algorithm and the optimal number of neurons in the hidden layer for predicting effective power, the authors of the given text used both the scaled conjugate gradient (SCG) and Levenberg-Marquardt (LM) learning algorithms with different numbers of neurons in the hidden layer. After testing different combinations, the authors found that the best learning algorithm and network architecture for predicting effective power was the LM: 5-6-1 architecture. This means that the network had 5 neurons in the input layer, 6 neurons in the hidden layer, and 1 neuron in the output layer.

	mpg	cylinders	horsepower	weight	acceleration	USA	Europe	Japan
mpg	1.000000	-0.777618	-0.778427	-0.832244	0.423329	-0.565161	0.244313	0.451454
cylinders	-0.777618	1.000000	0.842983	0.897527	-0.504683	0.610494	-0.352324	-0.404209
horsepower	-0.778427	0.842983	1.000000	0.864538	-0.689196	0.489625	-0.284948	-0.321936
weight	-0.832244	0.897527	0.864538	1.000000	-0.416839	0.600978	-0.293841	-0.447929
acceleration	0.423329	-0.504683	-0.689196	-0.416839	1.000000	-0.258224	0.208298	0.115020
USA	-0.565161	0.610494	0.489625	0.600978	-0.258224	1.000000	-0.591434	-0.648583
Europe	0.244313	-0.352324	-0.284948	-0.293841	0.208298	-0.591434	1.000000	-0.230157
Japan	0.451454	-0.404209	-0.321936	-0.447929	0.115020	-0.648583	-0.230157	1.000000

Fig 10: Correlation between different attributes

The authors also found the best learning algorithms and ANN architectures for other output parameters, such as average effective pressure.

3.4 Training and Testing

he choice of the training and testing data percentages is a critical aspect when constructing an Artificial Neural Network (ANN) architecture. A review of the existing literature indicates the use of different ratios for training and testing data [26-29]. These ratios commonly range from 90% training and 10% testing [23, 30], 85% training and 15% testing [27], 80% training and 20% testing [17], 75% training and 25% testing [25], to 70% training and 30% testing [24, 31].

In this specific study, 55 sets of experimental results were generated for the purpose of training and testing the ANN. The decision was made to allocate **80% of the data for training and 20% for testing**. Consequently, 44 data points were designated for training, while 11 data points were reserved for testing, selected in a random manner.

3.5. Normalization of data

IT administrators help determine capacity thresholds for apps in the private cloud. When workload capacity nears its threshold, the used application automatically switches over into the public cloud and traffic is pointed towards it. As the spike in resource needs diminish, the application is relocated back to the private cloud or on-premises infrastructure.
In essence, data normalization is the act of “cleaning” the data in order to make data input more coherent. A data collection is reorganized to eliminate any duplicate or unstructured information in order to make room for a better, more sensible way to store the data. This process is known as normalization.
A uniform data format for your who”e sy’tem Is the primary objective of data normalization. This makes it easier to query and evaluate the data, which can help in making better business decisions.

3.6 Statistical Evaluation

Ramp descent is the algorithm used in the BP training process. By adjusting the weights via its ramp, the BP algorithm seeks to reduce overall error and enhance network performance.
When the tested RMS values start to rise instead than fall, the network’s training is terminated. At this point, the findings are verified using testing data that has never been given to the network to see whether an ANN is producing accurate predictions [33]. For comparisons, the statistical techniques of RMS, R2, and MEP values have been applied.

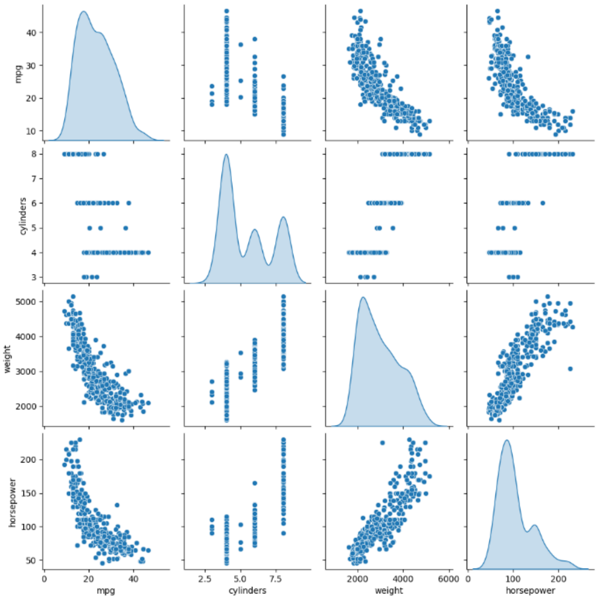


Fig 11: Pairplot of data

4. Results and Discussion

The efficient constant engine speed, pressure rises proportionately with increasing torque (Fig. 3a). Engine output rises in tandem with the motor speed at steady

torque levels. This conduct is the typical a feature of engines with internal combustion. Based on the engine speed, but at low torque values, the rate of power gain is minimal. At high torque values, it is discovered to be higher. These two situations may be accounted for by the lowest pumping and heat-loss values at higher engine speed, the ideal turbulence value, the maximum temperature and pressure within the cylinder, the fuel-to-air ratio consistently being at the same level throughout the combustion chamber, as well as the rise in filling.
This is caused by the abundant oxygen in the methanol, the combustion chamber's optimal conditions for this rate of burning, and the high thermal efficiency resulting from the suction pad's low temperature and high density. Because it is directly related to torque, as shown in Fig. 3b, the average effective pressure increases linearly. An increase in torque results in an improvement in thermal efficiency, which raises the average effective pressure.

The group known as Tensorflow Sequential holds the stack of linear format made up of several layers from the tf.keras.Model library package. The Module, Layer, and Model classes are the sources of this Sequential class's inheritance. Providing conclusions and training the module are the fundamental functions of Sequential. We shall examine what TensorFlow sequential is, its model, its function, its methods, its examples, and its methods of execution in this post before coming to a conclusion.

```
Model: "sequential"
_________________________________________________________________
 Layer (type)                Output Shape              Param #
=================================================================
 dense (Dense)               (None, 7)                 56

 dense_1 (Dense)             (None, 249)               1992

 dense_2 (Dense)             (None, 249)               62250

 dense_3 (Dense)             (None, 149)               37250

 dense_4 (Dense)             (None, 1)                 150

=================================================================
Total params: 101698 (397.26 KB)
Trainable params: 101698 (397.26 KB)
Non-trainable params: 0 (0.00 Byte)
_________________________________________________________________
```

Fig 12: Details of layers in our Network

As you train the deep learning optimizer model, adjust the weights for each epoch and reduce the loss function. An optimizer is a function or algorithm that modifies the neural network's parameters, such learning rates and weights. As a result, it aids in raising accuracy and decreasing total loss. With millions of parameters in most deep learning models, selecting the appropriate weights for the model is a difficult issue. It highlights the necessity of selecting an appropriate optimization algorithm for your use case. Therefore, before delving deeper into the topic, data scientists must comprehend these machine learning techniques.

5. Conclusion

The performance, average effective pressure, and exhaust gas temperature parameters of a methanol engine have all been studied in relation to the application of ANNs. N, T, m, Tin, and Tcw are the inputs used to train the network, and BSFC, Tex, Ape, and Pe are the outputs. This research uses artificial neural network (ANN) modelling of a methanol engine to forecast the engine's average effective pressure, exhaust gas temperature, effective power, and brake-specific fuel consumption. An ANN model for the engine based on a typical back propagation approach was created using a portion of the experimental data for training.
Next, by contrasting the predictions with the experimental findings that weren't included in the training process, the effectiveness of the ANN forecasts was evaluated.
The R2 score, also known as the coefficient of determination, is a statistical measure that evaluates the performance of a linear regression model. It determines the proportion of variance in the dependent variable that can be explained by the independent variable. In other words, it shows how well the data fit the regression model.

$$R^2 = 1 - \frac{SS_{RES}}{SS_{TOT}} = 1 - \frac{\sum_i (y_i - \hat{y}_i)^2}{\sum_i (y_i - \overline{y})^2}$$

The input layer was comprised of engine speed, torque, fuel flow, intake manifold mean temperature, and cooling water entry temperature; the output layer was composed of brake-specific fuel consumption, engine power average effective pressure, and exhaust manifold temperature. Following training, it was discovered that the training and testing data's R2 values are very near to For the testing data, the RMS values are less than 0.015 and the mean errors are less than 3.8%. It would be simple to conclude that the findings fall within the permissible bounds.
Our model predicts at the rate of **76.05 %** accuracy.

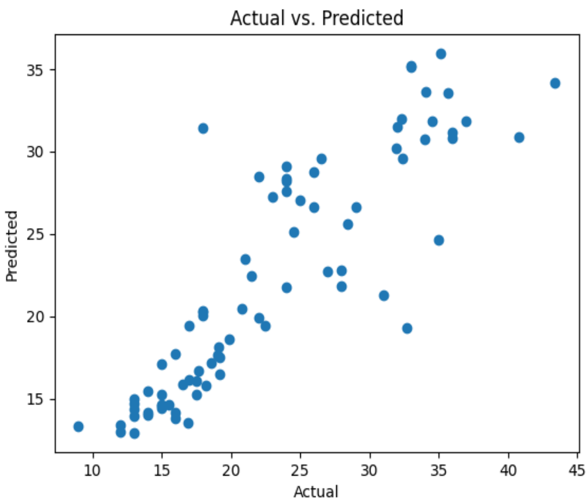


Fig 13: Actual V/S predicted value

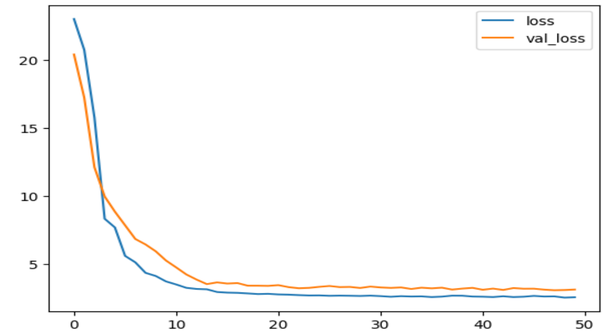


Fig 14: Loss V/S Validation loss

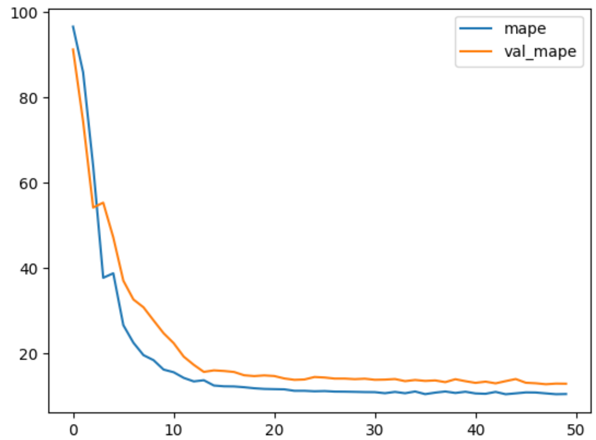


Fig 15: MAPE V/S Validation MAPE (Mean Absolute Percentage Error)

```
r2_score(y_test,y_pred)

0.7605934392603437
```

Fig 16: Final Result

6. Acknowledgement

I would like to express our deepest gratitude to all those who have contributed to the successful completion of this term paper.

First and foremost, I would like to thank our teacher for providing us with invaluable guidance and support throughout the entire process. Your expertise and insights have been invaluable in shaping the direction and focus of this work. Our gratitude also goes to the various scholars and researchers whose published work has informed our research, and to the various institutions and organizations who have provided us with access to data and resources. Finally, I would like to thank our friends and family for their constant encouragement and support. Your love and understanding have sustained us through the challenges of this project and beyond. Thank you all for your contributions, both large and small, towards the successful completion of this paper.

14. References

1. [1] A. Surmen, M.I. Karamangil, R. Arslan, Engine Thermodynamics. Aktuel Publishing, Istanbul, 2004.
2. [2] A. Calisir, M. Gumus, The effect of gasoline-methanol blends on the engine performance and exhaust emission in a spark ignition engine. IATS 2009: Proceedings of the 5th International Advanced Technologies Symposium; 2009 May 13-15; Karabük, Turkey. p. 189e1898.
3. [3] M.B. Celik, B. Ozdalyan, F. Alkan, The use of pure methanol as fuel at high compression ratio in a single cylinder gasoline engine, Fuel 90 (2011) 1591e1598.
4. [4] J. Li, C.M. Gong, Y. Su, H.L. Dou, X.J. Liu, Effect of injection and ignition timings on performance and emissions from a spark-ignition engine fueled with methanol, Fuel 89 (2010) 3919e3925.
5. [5] C. Sayin, Engine performance and exhaust gas emissions of methanol and ethanolediesel blends, Fuel 89 (2010) 3410e3415.
6. [6] E. Oztemel, Artificial Neural Network. Papatya Publishing, Istanbul, 2003.
7. [7] G. Zhang, B.E. Patuwo, M.Y. Hu, Forecasting with artificial neural networks: the state of the art, International Journal of Forecasting 14 (1998) 35e62.
8. [8] A. Negarestani, S. Setayeshi, M. Ghannadi-Maragheh, B. Akashe, Estimation of the radon concentration in soil related to the environmental parameters by APe
9. [9] M. Taheri, A. Mohebbi, Design of artificial neural networks using a genetic algorithm to predict collection efficiency in venturi scrubbers, Journal of Hazardous Materials 157 (2008) 122e129.
10. [10] I. Korkut, A. Acir, M. Boy, Application of regression and artificial neural network analysis in modeling of tool-chip interface temperature in machining, Expert Systems with Applications (2011). doi:10.1016/j.eswa.2011.03.044.
11. [11] A. Kurt, Modelling of the cutting tool stresses in machining of Inconel 718 using artificial neural networks, Expert Systems with Applications 36 (2009) 9645e9657.
12. [12] J. Porteiro, J. Collazo, D. Patiño, J.L. Míguez, Diesel engine condition monitoring using a multi-net neural network system with nonintrusive sensors, Applied Thermal Engineering 31 (2011) 4097e4105.
13. [13] T. Boushaki, S. Guessasma, J.C. Sautet, Predictive analysis of combined burner parameter

effects on oxy-fuel flames, Applied Thermal Engineering 31 (2011) 202e212.
14. [14] M. Khandelwal, T.N. Singh, Prediction of macerals contents of Indian coals from proximate and ultimate analyses using artificial neural networks, Fuel 89 (2010) 1101e1109.
15. [15] M. Rajendra, P.C. Jena, H. Raheman, Prediction of optimized pretreatment process parameters for biodiesel production using ANN and GA, Fuel 88 (2009) 868e875.
16. [16] R.M. Balabin, E.I. Lomakina, R.Z. Safieva, Neural network (ANN) approach to biodiesel analysis: analysis of biodiesel density, kinematic viscosity, methanol and water contents using near infrared (NIR) spectroscopy, Fuel 90 (2011) 2007e2015.
17. [17] B. Ghobadian, H. Rahimi, A.M. Nikbakht, G. Najafi, T.F. Yusaf, Diesel engine performance and exhaust emission analysis using waste cooking biodiesel fuel with an artificial neural network, Renewable Energy 34 (2009) 976e982.