Materials Today: Proceedings 38 (2021) 280–284

Contents lists available at ScienceDirect

Materials Today: Proceedings

journal homepage: www.elsevier.com/locate/matpr

Predictive modeling of engine emissions using machine learning: A review

Shivansh Khurana *, Shubham Saxena, Sanyam Jain, Ankur Dixit

Department Of Mechanical Engineering, ABES Engineering College, Ghaziabad, India

ARTICLE INFO

Article history:
Received 2 May 2020
Received in revised form 5 July 2020
Accepted 10 July 2020
Available online 12 August 2020

Keywords:
Predictive modelling
Engine emissions
Bio-fuels
Machine learning
Artificial intelligence

ABSTRACT

The increase in population has led to increase in demand of automobiles across the globe. Concerns about the pollutants emitted from an engine are growing periodically. The paper has tried to show an exploratory review of the various methodologies adopted for accurately measuring and analysing exhaust engine emissions. The paper has the objective of showing the main conclusions of the recent research performed since last decade (2008 to 2020) for the above mentioned topic with the help of artificial intelligence methodologies. This review addresses an important application of artificial intelligence that can be applied in measuring the engine emissions. The measurement of these engine emissions are mandatory for every automobile company. This task is usually achieved by repeated testing of automobile which is not a cost efficient process. These process usually involves expensive test rigs installation. But predictive modeling for accurate testing of emissions can be used as a digital/virtual tool. Hence there is an extensive scope of research in this field. The paper has tried to showcase emerging trend of advanced machine learning algorithms that can easily help in generating emission data in a simple manner.

Selection and Peer-review under responsibility of the scientific committee of the 2nd International Conference on Future Learning Aspects of Mechanical Engineering.

1. Introduction

Fuel has always been considered as an imperative parameter for driving the world's economy. But due to heavy dependence of various sectors especially transportation sector upon crude oil and its derivatives, we are on the verge of seeing their depletion soon. This also draws attention towards the negative impacts of this dependence like pollution, global warming due to the engine emissions of the fuel. Therefore, it has become necessary to create an effective system to measure these harmful emissions. To tackle this problem, initiatives are taken to develop alternative sources of fuels. In recent trends, research on Bio-fuels has significantly risen because of their myriad benefits. Reduced emission levels and easy manufacturing are other benefits to consider. Various approaches are used to increase the efficiency of Bio-Fuels and to generate optimum blend ratio with diesel and petrol. And for selectivity of best fuel blend, predictive modelling of engine emissions is taken into consideration. Predictive modeling is a process of generating outcomes using a statistical approach and concepts of probability. This approach requires certain input data to show accurate results. More the experimental data fed to the model, more is the accuracy of that prediction. Exactness of prediction is the key element which decides how accurate the model is. In this paper, we have also discussed about new technological advancements in the field of Machine Learning which also employs along with Artificial Intelligence to generate a highly exact and compact model.

2. Review

In 2008, Karri V. et al. [1] have identified various AI models as virtual sensors to investigate and measure all the possible engine emissions produced from a hydrogen powered car. The identification of these emissions are done using advanced artificial intelligence models. These advanced techniques includes adaptive neuro fuzzy approach, back-propagation with Levenberg–Marquardt algorithm, and an in-house software built by University of Tasmania. The research has included many important input parameters that has direct impact over the emissions produced by hydrogen powered engine [2] (Fig. 1). The experimentation process is highly sophisticated and has many advanced probes and sensors. The use of such sensors was also necessary in order to cap-

* Corresponding author.

https://doi.org/10.1016/j.matpr.2020.07.204
2214-7853/
Selection and Peer-review under responsibility of the scientific committee of the 2nd International Conference on Future Learning Aspects of Mechanical Engineering.

Fig. 1. Showing general working of adopted methodology.

ture the emission readings as accurately as possible. The input parameters that were included, is shown in below table (Table 1).

The output variables included CO, CO_2, HC, $NO_{x..}$ The use of these advanced artificial intelligence methodologies are termed as "virtual sensors" in the mentioned paper. The predictions obtained were quantitative in nature. The accuracy of the models adopted came out to be 95% accurate when compared with actual experiment data [3] (Fig. 2).

In 2009, Ghobadian B. et al., [4] created artificial neural network model to predict brake power, torque, and engine emissions as primary output. The model was tested on diesel engine which was fueled by bio-diesel produced from waste vegetable cooking oil. The process of experimentation was started by manufacturing bio-fuel after the collection of vegetable waste cooking oil. Then this manufactured bio-fuel was fed into a two-cylinder, water cooled diesel engine and engine emissions along with all the mentioned performance parameter readings were observed and calculated. Various blends of bio-fuels were made and each blend has undergone same testing procedure. By adopting this methodology, significant amount of data was produced which helped ANN model to train itself. The results after creating and testing of the results were done, showed that back-propagation algorithm was enough in giving accurate predictions of engine torque, brake power and emissions. The mean squared error calculated was 0.0004 which is quite less and is tending towards ideal value.

Table 1
Input parameters.

RPM
Throttle
Air flow mass
Manifold pressure
Power
Engine temp.
Air temp.
Exhaust gases tem.

In 2010, again a similar approach was adopted by Kiani M. et al. [5] ANN model was applied in predicting engine performance and engine emissions. The main difference was the type of engine used. The engine used was 4 cylinder, four stroke spark ignition engine and hence it can be shown that these artificial models are very much adaptable when it comes to their application of use. The fuel used to run this engine was ethanol gasoline based. The results produced by this model were again nearly accurate and fast.

In 2011, Gopalakrishnan K. et al. [6] identified the problem of rise of carbon emissions from passenger vehicles. It is an undeni-

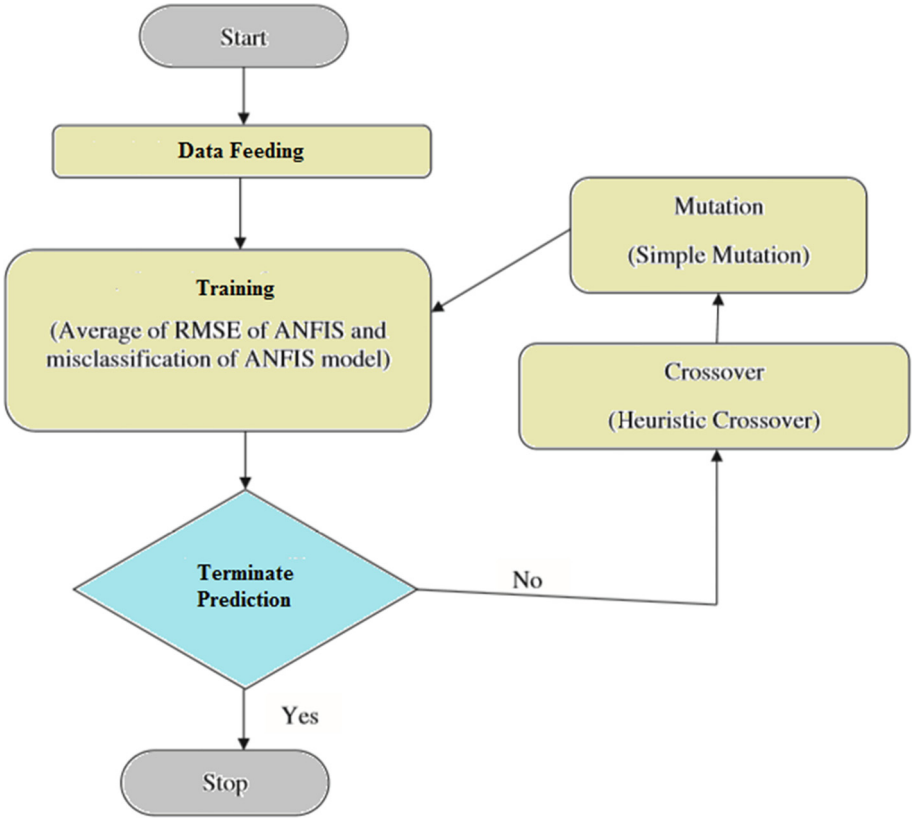

Fig. 2. Showing Optimization of ANFIS model when implemented with GA.

able fact that public transit buses are the most common mode of transportation for general public. The study showed the application of neuro-fuzzy model in order to predict the emissions generated by actual biodiesel powered transit buses. The data generation part was done using state of the art device "Portable emissions measurement system". The neuro-fuzzy techniques are quite capable for in order to handle noise in the data, this is necessary in order to handle any possibility for over fitting. The two neuro-fuzzy techniques that are used in this experiment are: - Adaptive Neuro-Fuzzy Inference System (ANFIS) and the Dynamic Evolving Neuro-Fuzzy Inference System (DENFIS).

Although the ANFIS model showed better accuracy in regards of the whole study. The pollutants that are under study in this paper are (NOx, HC, CO, CO2 and PM). The two fuels that have been used in this study are B10 and B20. According to various research use of biodiesel instead of diesel decreases HC, CO, CO2 and PM, but increases NOx emissions. Various Researches have shown that bio fuels are non-hazardous, biodegradable, nontoxic, having high lubricity and high flash point. The primary reason why the bio fuels are coming into limelight is that it increases the lubricant properties of the fuel it's mixed with. Hence the emissions from PEMS used in transit bus powered by variety of biodiesel blends were measured and used in developing emissions prediction models using the two models as already been discussed. Now since ANFIS gave only one output, five models were created in order to get the predictions for all the five different types of emissions. The predictions obtained by the ANFIS model were quite accurate for NOx and CO2. The accuracy for the ANFIS model is determined by the R2 value.

Herein the data was studied in the real world scenario using the neuro-fuzzy techniques. However the data was obtained using PEMS. Both ANFIS and DENFIS models showed highest accuracies for NOx and CO2 emissions.

In 2012, J. Mohammadhassani et al. [7] created a model for prediction of NO_x emissions in direct injection engine. The approach used was ANN technique. As stated by the research paper, Direct Injection system is a propulsion system for low fuel consumption and large power output. This system is mainly used in the automotive industry. Many studies have been done on NOx emissions and it has various environmental damaging effects [4]. The diesel engine industry has encountered huge technological advancement creating various devices like Electronic Control unit and Engineering Management System. It provides insights on how to optimize the best suited features in order to enhance the performance. With the help of above mentioned paper, Artificial Neural technique is used to create the relationship between NOx emissions and operating parameters of test engine.

The operating parameters used for testing were such as opening of throttle valve, fuel Injection timing, speed in rpm and consumption of fuel. Experimental tests were performed on 144 speeds varied between 591 and 2308 rpm. The LM training algorithm is also used for training the system. A MATLAB program was also created to identify the correlations between validation and testing. About 90% of data set is used for training purpose. The results produced by ANN approach were validated through experimental results and showed a great deal of correlation (about 0.92 for training, 0.98 for validating and 0.89 for testing) with the observed data. Hence ANN model was successfully deployed and showed accurate results.

In 2013, [8] similar approach was adopted and same technique was followed. The approach showed output for turbo charged engine was predicted which included its performance analysis as well as engine emissions. The paper is given provide in the reference section. In 2014, [9] the same ANN model was proposed and applied. The main difference was the use of a CNG dual fuelled diesel engine. It predicted brake specific fuel consumption, emissions and performance parameters. Yet again ANN showed accurate results.

In 2015, [10] ADARSH A.R. et al, showed the use of other advanced artificial intelligence models in this field. The research paper showed prediction modeling of engine performance using various techniques like, ANN, Genetic Algorithm (GA) and ANFIS. The model was firstly created using ANFIS model and accuracy was calculated. Then GA was implemented to optimize ANFIS model and results showed that the previous model's accuracy was significantly increased, which is 46.6%.

In 2016, Deniz S. et al. [11] published research paper which showed the scope of data mining in the field of testing engine performances and curbing emissions. The paper showed various data mining techniques such as neural networks, Bayesian networks, C5.0 algorithm. The scope of the paper showed that such artificial methodologies can help in creating efficient and affordable automobiles which are also eco-friendly in nature. The work was divided into several phases. These phases are shown in Fig. 3.

The paper showed how data mining techniques can successfully be used in order to solve the emerging problems in automobile industry. The data used for these models was based on real time values and scope was fully practical.

In the similar year, Li Q. et al. [12] used machine learning approach on idling emissions of light duty vehicles. There have been many cases when a car or any other road vehicle has to stop either voluntarily or due to some kind of traffic jams. The emissions produced during the resting period of vehicle but with engine running are also enormous. Idling refers to a condition in which the engine of a car is running but the vehicle is not in motion.

There are two types of idling, Discretionary idling and non-discretionary idling. Discretionary idling is defined as the state in which the driver decides to put the vehicle at rest. While non-discretionary idling is when the driver has to stop the vehicle due to certain conditions like traffic light, traffic jams etc. The above mentioned paper shows the emission modeling of engine at idling condition only. The paper implemented machine learning techniques: KNN (K- nearest neighbour, Bagged Decision Trees, Neural Network, Support Vector Machine, CHAID model).

Attempts have been made to find out the best suited algorithms for best fit models based on real time driving tests. The accuracy of these models is compared by known metrics of statistics such coefficient R and RMSE value.

Out of the five algorithms, two models- KNN and BBDT models were selected since it has been found that these two models show very less relative errors among five of them. Then after choosing these two models, their outcomes were compared by their root mean square error. BBDT algorithm showed very less deviation from the RMSE.

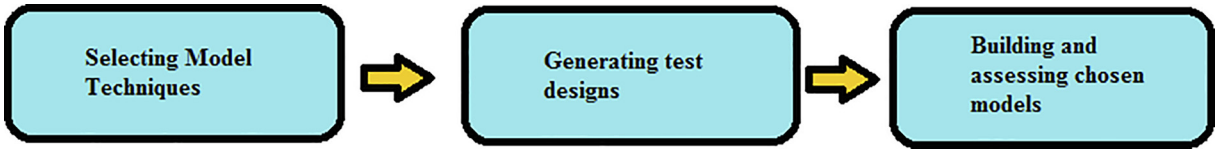

Fig. 3. General Phases adopted.

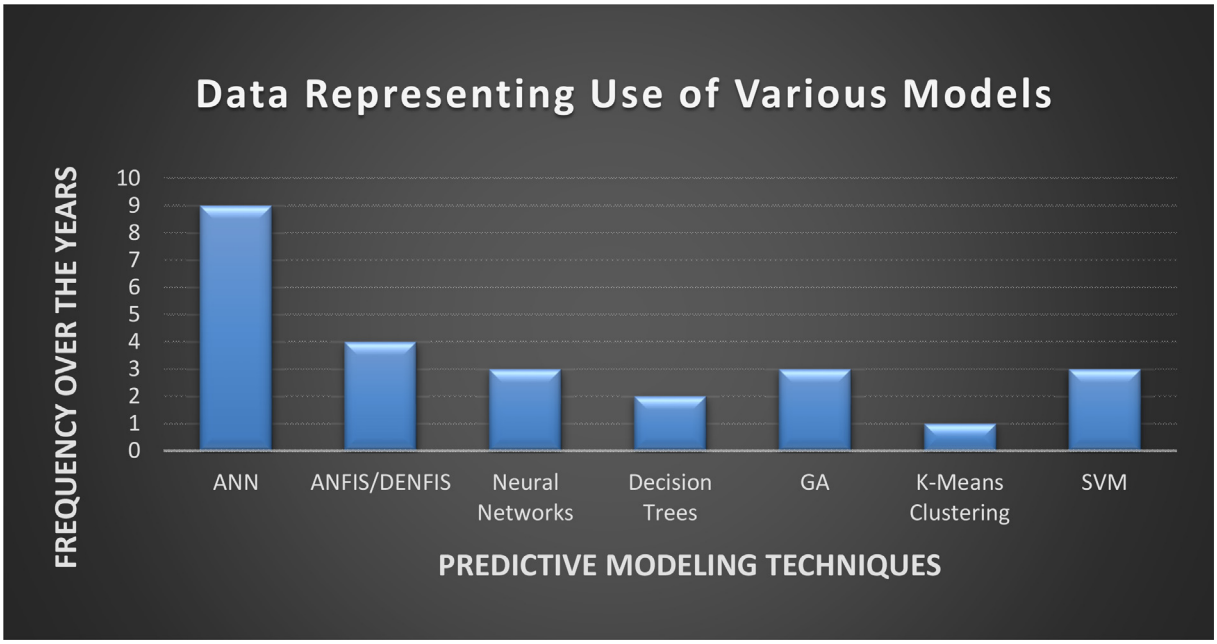

Fig. 4. Outcome.

In 2017, Dharma S. et al. [13] has done ANN modeling on Jatropha based biodiesel for predicting performance and emission of diesel engine. ANN model was implemented using back propagation technique. Yet again ANN worked accurately by showing 98% coefficient of determination (R^2) value. The parameters that are analysed are: Brake thermal efficiency, brake power, and exhaust emissions.

In 2018, Lukose A. et al, [14] proposed a way to control air–fuel ratio of diesel engine even in transient conditions. Exhaust gas emissions such as CO2, NOx, PM, CO and unburned HC were inspected and monitored closely. For every input parameter, corresponding emission was produced using Genetic Algorithm. The paper showed the versatile nature of genetic algorithm. Since this algorithm produces output through natural selection, the paper showed that how it can used for a problem in which information is limited. The factors considered for analysing output parameters were under transient condition. The values which were under observation were plotted against the values which were predecided. From this method Lagranges interpolation equation was calculated and fed to GA for optimization process.

In 2019, Azeez O.S. et al., [15] used optimized ANN model to predict CO emissions and also generate daily localized maps of the specified place. The emissions recorded were produced by traffic vehicles of that specified place in the research paper. The paper represented hybrid models of data mining and Geo spatial information system (GIS) which has the motto of design, analyse, and store the different types of spatial data. The ANN model used was modified with correlation based feature selection model to generate CO emissions constantly throughout the day (Four times a day). Then to generate prediction maps so that level of CO emissions can be analysed at a localized place. The maps generated through this model showed the concentration of CO emissions produced from traffic vehicles. The model used six parameters: number of vehicles, number of heavy vehicles, number of motorbikes, temperature, wind speed and a digital surface model. It has been stated that 86% accuracy was achieved in this model. This model can be of great use when it comes to monitoring localized prediction of emission that too from daily traffic vehicles. The model can be very helpful for knowing the patterns of harmful emissions and at what time these emissions are at peak. Lastly, in 2020 [16], a recent research paper adopted ANN approach for performance evaluation of SI engine. The paper stated that isoamyl alcohol group based fuel has less information regarding its use. The experiments were performed using test rigs for creating emission data. The dataset was created by varying certain input parameters like opening of throttle valve, varying engine speeds and changing compression ratio. The data was then fed to ANN model. The ANN model was further optimized with RSM (Response Surface Methodology). The correlation factor was found ranging in 0.94 to 0.99.

3. Outcome

From observing the statistical data it can be seen that Artificial Neural Network is widely used when it comes to predictive modeling of engine emissions. This has open up new area of research field to find more and more optimum algorithms and advanced techniques to increase the accuracy of the results (Fig. 4).

4. Conclusion

This paper has reviewed the literature of the research papers that included real time emission measurement and validating those measured values with machine learning algorithm outputs. This has opened up an interesting discussion for how artificial intelligence can be useful in predicting the emission rates of exhaust gases of an automobile. From the data collected after reviewing these research papers, it is evident that ANN (Artificial Neural Network) has been an accurate and widely used algorithm in addressing this kind of problem. Since it is highly adaptable, it can be easily optimized with other algorithms making it more robust.

The traditional way of analyzing and measuring an automobile emissions typically involves two ways: Laboratory testing and Field testing. These type of testing are normally highly sophisticated and expensive in nature. They usually involve lots of expensive equipment. But implementing any of the technology above

can make these testing efficient since they improve more and more when large amount of data fed to them is increased. It is evident from these literatures that once the process of training the model is over, they can be instantly implemented for practical use and show accurate results. Hence application of any of the methodologies discussed in this paper can easily make these testing simple and cost effective in nature.

CRediT authorship contribution statement

Shivansh Khurana: Conceptualization, Methodology, Software, Formal analysis, Investigation, Writing - original draft, Writing - review & editing, Visualization. **Shubham Saxena:** Validation, Formal analysis, Data curation, Writing - original draft, Writing - review & editing. **Sanyam Jain:** Software, Data curation, Writing - original draft, Writing - review & editing, Visualization. **Ankur Dixit:** Supervision, Project administration, Conceptualization.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] V. Karri, T. Ho, Predictive models for emission of hydrogen powered car using various artificial intelligent tools, Neural Comput. App. 18 (2009) 469–476.
[2] H. Hafez, V. Karri, M. Kristiansen, Preliminary investigation to use bayesian networks in predicting NOx, CO, CO2 and HC emissions. In: Proceedings of IASTED international conference on energy and power systems. Krabi, Thailand, 2005.
[3] M. Negnevitsky, Artificial Intelligence, A guide to intelligent systems, Addison-Wesley, Harlow, 346–349, 2005.
[4] B. Ghobadian, H. Rahimi, A.M. Nikbakht, T.F. Yusaf, Diesel engine performance and exhaust emission analysis using waste cooking biodiesel fuel with an artificial neural network, Renew. Energy. 34 (2009) 976–982.
[5] S.H. Hosseini, A.T. Ahmad, B. Ghobadian, A.M. Ahmad, Application of an artificial neural network model for prediction of diesel engine heat using nano-additives in diesel-biodiesel blends, Agri. Eng. Int.: The CIGR e-Journal 19 (2017) 76–83.
[6] K. Gopalakrishnan, A. Mudgal, S. Hallmark, Neuro-fuzzy approach to predictive modeling of emissions from biodiesel powered transit buses, Transport. 26 (2011) 344–352.
[7] J. Mohammadhassani, S. Khalilarya, M. Solimanpur, A. Dadvand, Prediction of NOx Emissions from a Direct Injection Diesel Engine Using Artificial Neural Network. Modelling and Simulation in Engineering. Vol 2012, 2012.
[8] O. Özener, L. Yüksek, M. Özkan, Artificial neural network approach to predicting engine-out emissions and performance parameters of a turbo charged diesel engine, Thermal Sci. 17 (2013) 153–166.
[9] S. Roy, R. Banerjee, A. Das, P. Bose, Development of an ANN based system identification tool to estimate the performance-emission characteristics of a CRDI assisted CNG dual fuel diesel engine, J. Natural Gas Sci. Eng. 21 (2014) 147–158.
[10] A. Rai, P. Pai, B. Rao, Prediction models for performance and emissions of a dual fuel CI engine using ANFIS, Sadhana. 40 (2015) 515–535.
[11] S. Deniz, H. Gökçen, G. Nakhaeizadeh, Application of data mining methods for analyzing of the fuel consumption and emission levels, Int. J. Eng. Sci. Res. Technol. 5 (2016) 377–389.
[12] Q. Li, F. Qiao, L. Yu, A machine learning approach for light-duty vehicle idling emission estimation based on real driving and environmental information, Environ. Pollution and Climate Change 1 (2016) 1000106.
[13] S. Dharma, H.H. Masjuki, H.C. Ong, A. Sebayang, F. Kusumo, J. Milano, Experimental study and prediction of the performance and exhaust emissions of mixed Jatropha curcas-Ceiba pentandra biodiesel blends in diesel engine using artificial neural networks, J. Cleaner Prod. 164 (2017).
[14] A. Lukose, A.V.N. Krishna, Performance evaluation of diesel engine by feature selection of exhaust using machine learning algorithm, Int. J. Eng. Sci. Invention (IJESI). 7 (2018) 24–28.
[15] O. Azeez, B. Pradhan, H. Shafri, N. Shukla, H. Rizeei, Modeling of CO emissions from traffic vehicles using artificial neural networks, Appl. Sci. 9 (2019) 313.
[16] S. Uslu, M. Celik, Performance and exhaust emission prediction of a SI engine fueled with I- amyl alcohol-gasoline blends: An ANN coupled RSM based optimization, Fuel 265 (2019).