Vol. 02, No. 2,(2018), 049–057 , ***ISSN: 2581-3242***

IJMLNCE JOURNAL

International Journal of Machine Learning and Networked Collaborative Engineering

Journal Homepage: http://www.mlnce.net/home/index.html

DOI : https://doi.org/10.30991/IJMLNCE.2018v02i02.002

Prediction Model for Pollutants with Onboard Diagnostic Sensors in Vehicles

[a]Bryce Maldonado[*], [b]Malik Bennabi, [c]Vicente García-Díaz, [d]Cristian González García and [e]Edward Rolando Núñez-Valdez

[a]*University of South Florida, Florida, USA*

maldonado@mail.usf.edu

[b]*University of Sakarya, Faculty of Computer and Information Sciences, Sakarya, Turkey*

malik.bennabi@ogr.sakarya.edu.tr

[c,d,e]*Department of Computer Science, University of Oviedo, Spain*

[c]*garciavicente@uniovi.es,* [d]*gonzalezcristian@uniovi.es,* [e]*nunezedward@uniovi.es*

Abstract

In this work, a prediction model is developed to illustrate the relationship between the internal parameters of a vehicle and its emissions. Vehicles emit various hazardous pollutants and understanding the influence of in-vehicle parameters is key to reducing their environmental impact. The values of the internal parameters were collected through the On-Board Diagnostics port, while the values of the emissions were measured from the exhaust pipe using Arduino sensors. The observed values were then matched based on the timestamps received from both sources and fit with both linear and polynomial regressions to accurately model the relationship between the internal parameters and pollutants. These models can then be used to estimate vehicle emissions based on the in-vehicle parameters, including vehicle speed, relative throttle position, and engine revolutions per minute. A wide majority of the relationships between various in-vehicle parameters and emissions show no observable correlation. There are observable correlations between carbon dioxide emissions and vehicle speed, as well as carbon dioxide emissions and engine revolutions per minute. These relationships were modelled using linear and polynomial regression with a resulting adjusted R-squared value of approximately 0.1.

Keywords

41A05,

41A10,

65D05,

65D17,

Artificial Intelligent,

Prediction model,

Sensors,

OBD

1. Introduction

The concentration of car emissions is a crucial factor for manufacturers to consider when designing and manufacturing vehicles. The burning of fossil fuels to power these automobiles creates various

[*] Corresponding author

Bryce Maldonado

Email: maldonado@mail.usf.edu

emissions, some of which are toxic and hazardous to human health [1,2]. Standards and regulations have been developed and refined over the past two decades to address various health and environmental concerns, leading to the creation of institutions and agencies such as the Environmental Protection Agency (EPA) in the United States and the European Environment Agency (EEA) in Europe, to ensure that manufacturers adhere to certain protocols.

Reducing emissions goes beyond the manufacturer and the composition of the fuel used [3], as consumers can participate in eco-driving to further reduce emissions. Various eco-driving programs provide tips and guidelines on methods and techniques to further reduce vehicle emissions by 10% on average [4]. These eco-driving programs emphasize the importance of anticipating traffic, removing unneeded heavy loads from the vehicle, and other small behavioral changes that can reduce emissions [5].

There are many driving factors that influence the fuel consumption of a vehicle, and therefore the emissions, but the plausibility and viability of these eco-driving options may vary between individuals. Eco-driving strategies are being tested and implemented to assist drivers in reducing their fuel consumption, especially in poor driving conditions including severely congested roadways [6]. To aid consumers desiring to participate in eco-driving, it is important to understand the correlations between various in-vehicle parameters and the emissions of a vehicle. Some internal parameters are affected by driving patterns and characteristics, and thus understanding the influence these parameters have on emissions can lead to new eco-driving strategies to be developed. The accessibility and ease of eco-driving can always be improved upon and attract and encourage drivers to reduce their fuel consumption in the long-run [5].

In this paper, various in-vehicle parameters of a car are observed and plotted against the outputted emissions from the exhaust pipe to determine the effect these parameters have on vehicle emissions. Any possible correlations between in-vehicle parameters and the concentrations of outputted emissions are analyzed through simple regressions [7]. This results in prediction models that can be used to determine the influential strength of various in-vehicle parameters on vehicle emissions.

The rest of this work is structured as follows: Section II examines the common pollutants in vehicle emissions and explores previous studies that have examined the significance of various vehicle and driving characteristics on emissions and fuel consumption. Section III describes the method and the reasoning behind the method for collecting the data. The implementation of the methodology is described afterwards. Section IV details the results and the analysis conducted on the recorded data. Section V presents the conclusions drawn from the study and Section VI explores ways and directions to further this work.

2. Background

Driving styles can be classified and deconstructed into identifiable characteristics to establish relationships between driving patterns and fuel consumption [8]. Internal car factors, such as acceleration, deceleration, average speed, RPM, and throttle position have been correlated directly to fuel consumption [8,9]. Managing steady non-volatile speeds and adhering to certain velocity ranges has yielded up to a 20% decrease in vehicle emissions in simulation, with significant reduction still observed in experimental cases 10. Sixty-two observable driving pattern parameters have been narrowed down to nine that have a significant effect on emissions 8. The parameters that have significant influence on emissions can be summarized by five key components: acceleration with high power demand, speed oscillation, extreme acceleration, deceleration, and stop 8. These components were used, in part, for determining which internal vehicle parameters to observe and plot with the emission concentrations when searching for relationships in the data because previous studies examine the effects of driving patterns on fuel consumption instead of emission concentrations directly.

There's been a demonstrated linear positive correlation between RPM and fuel consumption, and a polynomial positive correlation between relative throttle position and fuel consumption using statistical data analysis 9. This indirectly implies a relationship between these internal factors and emission concentrations through fuel consumption. The goal of this study is to expand upon this work to examine more than these two internal vehicle parameters, and to note their direct impact and emission concentrations. Past studies have examined the accuracy of On-Board Diagnostic systems using the observable concentrations of emissions from the exhaust pipe 11. Rather than comparing the observed concentrations of emissions to specific internal vehicle parameters, these outputted concentrations were classified by driving states (accelerating, cruise, idle) 11. This paper further seeks to examine the direct relationship between chosen internal vehicle factors and emission concentrations, modeling any observed correlations between these factors and emission concentrations to accurately estimate emission levels when driving.

Vehicle emissions mainly consist of nitrogen, carbon-dioxide, and water vapor 12; hazardous emissions such as carbon monoxide, nitrogen-oxide compounds, hydrocarbons, and particulate matter constitute less than 1% of the total emissions from motor vehicles 12. Most of the sensors for this study will examine the hazardous emissions, but a carbon dioxide sensor will also be utilized to examine a larger portion of emissions. A tool has been developed for simulating the concentrations of these emissions given various other internal parameters 13. The simulator uses a verified database of emissions concentrations to simulate emissions, instead of using real-time observed data to correlate the internal parameters to emission concentrations 13. A statistical model for vehicle emissions has been created previously 14, but not using only internal vehicle parameters and real-time observed data.

3. Experiment Description

The On-Board Diagnostics system was used to record the values of various internal car factors while driving. It provides data for many of the current conditions and parameters of a vehicle 15. The On-Board Diagnostics PIDs (OBD-II PIDs) were read using an OBD-port scanner while the vehicle was in use; Table 2 displays the recorded in-vehicle parameters obtained through the OBD-II port. Not all parameters are relevant, while the bolded parameters significantly related to fuel consumption and as such were the chosen to be analyzed. These parameters were chosen according to the key components of strong fuel consumption influence 8. To measure the emission concentrations, a set of MQ gas sensors were connected to an Arduino board, as seen in Figure 1; Table 2 provides the names of the sensors employed in the experiment and the gases they detected and measured. Automobiles produce a variety of emissions including carbon dioxide (CO_2), hydrocarbons (HC), nitrogen-oxide compounds (NO_x), and other particulate matter 12. Some MQ sensors used in the experiment detected gases that were not present or strongly represented, in standard vehicle emissions, such as alcohol; hydrocarbons and carbon monoxide make up less than 1% of the total emissions of a motor vehicle, while carbon dioxide constitutes a more observable 12% 12. Because of its observability, the MG811 sensor is the most significant sensor to monitor; the other hydrocarbon and carbon monoxide sensors were included to note any significant relationships between driving parameters and the more toxic vehicular emissions.

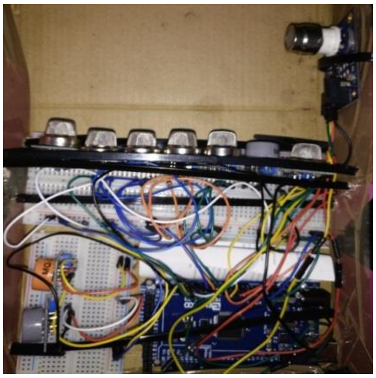

Fig.1. Arduino MQ2-9,135 & MG811 Sensors

Fig.2. Citroen C4 during data collection

4. Implementation

Three different automobiles were driven during the experiment to examine the influence in-vehicle parameters have on emissions regardless of the manufacturer and model of car; the chosen automobiles included a Citroen C4, a Citroen Xsara, and a Honda Accord. The vehicles tested in this experiment ran on diesel and gasoline. During the trial run with each vehicle, an Arduino wired with the set of MQ gas sensors was secured into a cardboard holding device and fastened to the end of the exhaust pipe, as seen in Figure 2. The automobile was driven in a nondiscriminatory manner around the city of Oviedo, Spain for an undetermined and variable length of time. Data recorded from both the OBD-II port and the MQ-sensors was collected from each vehicle and then synchronized through the timestamps included in the data. There is no fixing of Arduino values to account for possible delay between a change in an internal parameter and a change in the emissions out of the exhaust pipe. To justify this, a second pass between each full reading of the MQ sensors on the Arduino.

Table 1. Observed vehicle parameters

Internal Parameter	Units
Accelerometer(X)	m/s^2
Accelerometer(Y)	m/s^2
Accelerometer(Z)	m/s^2
Accelerometer(Total)	m/s^2
GPS Speed	m/s
OBD Speed	m/s
Intake Pressure(MAP)	Pa
RPM	rpm
Throttle Position	%
Relative Throttle Position	%

Table 2. Tested Sensors

Sensor	Gas(es) detected
mq2	Combustible gas, smoke
mq3	Alcohol
mq4	Methane, Propane, Butane
mq5	Butane, Propane, Methane
mq6	Liquefied petroleum, Butane, Propane, LPG
mq7	Carbon monoxide
mq8	Hydrogen
mq9	Carbon Monoxide, Methane
mq135	Ammonia sulfide, Benzene vapor
mg811	Carbon dioxide

5. Results

An overwhelming majority of the graphs displayed no significant observable relationship or correlation between an internal parameter and any of the MQ gas sensors. Some examples of these graphs are shown in Figures 3 and 4. All the parameters from Table 2 were examined, but only a few relationships produced graphs that demonstrated a perceivable correlation. These correlations can be seen in Figure 5 and 6 which clearly demonstrate a positive correlation between the independent and dependent variables, but the large variance of the data, especially at higher values, lessens the feasibility of an accurate estimation model. Regardless, a simple linear (shown in red) and polynomial regression (shown in blue) are run on each graph to model the general trend of the data. The resulting prediction model equations are given in Table 3, along with the adjusted R-squared value (ARV) and the residual standard error (RSE). The wide variance noted above in the plots has resulted in none of the adjusted R-squared values reaching over 0.2, meaning that the prediction models have an approximate accuracy of 17-20%. For the other emissions that the remaining sensors tested for, including CO, HC, and combustible smoke, there appeared to be no significant correlation between internal vehicle parameters and those emissions in this experiment.

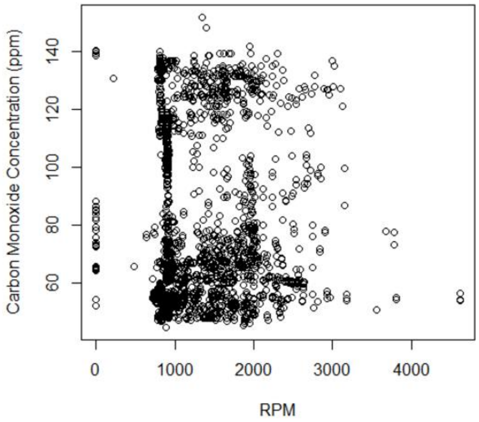

Fig. 3. RPM vs Carbon Monoxide Concentration

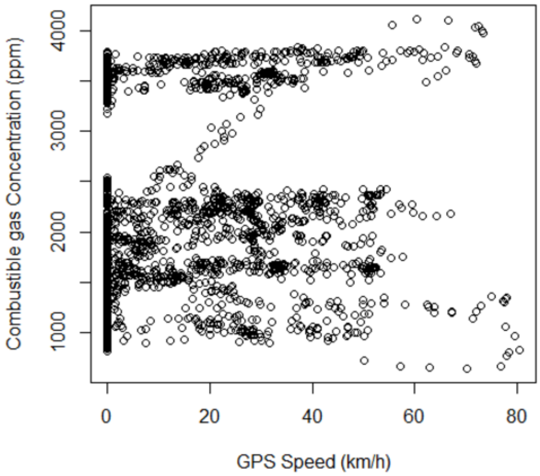

Fig. 4. GPS Speed vs Carbon Monoxide Concentration

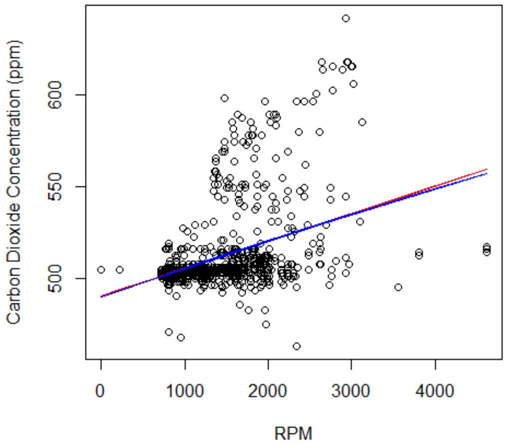

Fig. 5. RPM vs Carbon Dioxide Concentration

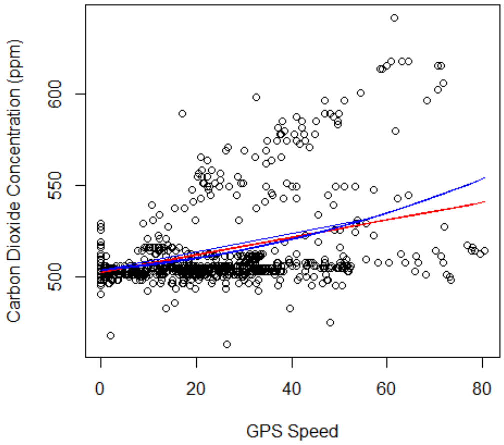

Fig. 6. GPS Speed vs Carbon Dioxide Concentration

6. Discussion

The lack of significant and accurate estimation results may stem from the method by which the data was collected. Having MQ gas sensors attached to the exhaust pipe of the vehicle may have led to volatile data collection since, as the vehicle moves, the emissions may move past the sensors before the MQ sensors have time to detect them. A total of 10 Arduino gas sensors were used during the experiment, and each sensor took 0.1 seconds to produce a new piece of data, meaning each sensor only took a reading from the exhaust pipe every second. This means that a vehicle moving at just 10 km/h would travel 2.77 meters before each sensor would read from the exhaust pipe. Additionally, while the Arduino sensors were secured on the opening of the exhaust pipe, it is possible for emissions to miss the sensors due to the nature of gases and the wind at high velocities. Gases expand to fit the size of their container, emissions may have expanded into the atmosphere once outside of the exhaust pipe and missed the MQ sensors.

Table 3. Regression Analysis

Regression type	Equation	ARV*	RSE#
Linear Regression RPM vs CO2	490.9+0.01484x	0.1762	18.11
Polynomial Regression RPM vs CO2	$490.3+0.01574x-2.663e{-}07x^2$	0.1756	18.12
Linear Regression GPS Speed vs CO2	502.788+0.47524x	0.1854	18.01
Polynomial Regression GPS Speed vs CO2	$504+0.2182x+0.00501x^2$	0.1942	17.92

*ARV–Adjusted R-squared Value
#RSE–False Residual Standard Error

7. Conclusions and Future work

Understanding the relationship between driving patterns and vehicle emissions is important when furthering eco-driving efforts. An estimation model for predicting and illustrating the relationships between internal vehicle parameters and emissions was sought in this experiment, to observe the significance of these parametersâ€™ influence on emissions. The data collected from the OBD-II port and the MQ sensors yielded significant relationships only between a couple internal parameters and the carbon dioxide emissions.

Hydrocarbons, carbon monoxide, and combustible smoke emissions were not significantly affected by variations in the recorded internal vehicle parameters.

There is a general positive correlation between engine RPM and carbon dioxide emissions, as well as vehicle speed and carbon dioxide emissions. This positive correlation has been fitted with estimation models from both linear and polynomial regressions. The resulting estimation models have adjusted R-squared values that are less than 0.2. While these estimation models do not accurately predict emission values, they do illustrate the positive relationship between the two variables, which is significant enough to be observed when plotted.

There exist multiple possible directions to continue the work of this experiment. Nitrogen oxide emissions constitute the largest proportion of vehicle emissions, this experiment could be repeated with nitrogen oxide and other hydrocarbon sensors to continue to search for significant relationships. Increasing the read time of the sensors and trapping emissions to collect data before releasing them into the air could provide more desirable results.

More tests can be run including a wider variety of vehicle makes and models. The vehicle selection can also be expanded by including hybrid vehicle, such as diesel-electric, and analyzing emissions with respect to in-vehicle parameters and fuel type.

8. Acknowledgements

This research has been partially supported by the National Science Foundation under grants No. 1458928 and No. 1645025, An REU Site on Ubiquitous Sensing.

The authors would like to thank the University of Oviedo, the University of South Florida, USA for organizing the research camp and providing us with all the support and assistance needed along this work.

The University of Sakarya of Turkey, for the indirect support of the participation of Malik Bennabi in this work.

All the professors who participated in the elaboration of the research camp content, activities, workshops and mentoring as well, particularly Dr. Miguel A. Labrador & Dr. Daniel F. Lanvin.

References

[1].Pearson RL, Wachtel H, Ebi KL. Distance-weighted traffic density in proximity to a home is a risk factor for leukemia and other childhood cancers. *J Air Waste Manage Assoc*. 2000;50(2):175-180. doi:https://doi.org/10.1080/10473289.2000.10463998.

[2]. Nicolai T, Carr D, Weiland SK, et al. Urban traffic and pollutant exposure related to respiratory outcomes and atopy in a large sample of children. *Eur Respir J*. 2003;21(6):956-963. doi:https://doi.org/10.1183/09031936.03.00041103a.

[3]. Zielinska B, Sagebiel J, McDonald JD, Whitney K, Lawson DR. Emission rates and comparative chemical composition from selected in-use diesel and gasoline-fueled vehicles. *J Air Waste Manage Assoc*. 2004;54(9):1138-1150. doi:https://doi.org/10.1080/10473289.2004.10470973.

[4]. Barkenbus JN. Eco-driving: An overlooked climate change initiative. *Energy Policy*. 2010;38(2):762-769. doi:https://doi.org/10.1016/j.enpol.2009.10.021.

[5]. Beusen B, Broekx S, Denys T, et al. Using on-board logging devices to study the longer-term impact of an eco-driving course. *Transp Res part D Transp Environ*. 2009;14(7):514-520. doi:https://doi.org/10.1016/j.trd.2009.05.009.

[6]. Saboohi Y, Farzaneh H. Model for developing an eco-driving strategy of a passenger vehicle based on the least fuel consumption. Appl Energy. 2009;86(10):1925-1932. doi:https://doi.org/10.1016/j.apenergy.2008.12.017.

[7]. Friedman J, Hastie T, Tibshirani R. *The Elements of Statistical Learning*. Vol 1. Springer series in statistics New York; 2001.

[8]. Ericsson E. Independent driving pattern factors and their influence on fuel-use and exhaust emission factors. *Transp Res Part D Transp Environ*. 2001;6(5):325-345. doi:https://doi.org/10.1016/S1361-9209(01)00003-7.

[9]. Lee MG, Park YK, Jung KK, Yoo JJ. Estimation of fuel consumption using in-vehicle parameters. *Int J u-and e-Service, Sci Technol*. 2011;4(4):37-46.

[10]. Barth M, Boriboonsomsin K. Energy and emissions impacts of a freeway-based dynamic eco-driving system. *Transp Res Part D Transp Environ*. 2009;14(6):400-410. doi:https://doi.org/10.1016/j.trd.2009.01.004.

[11]. Frey HC, Unal A, Rouphail NM, Colyar JD. On-road measurement of vehicle tailpipe emissions using a portable instrument. *J Air Waste Manage Assoc*. 2003;53(8):992-1002. doi:https://doi.org/10.1080/10473289.2003.10466245.

[12]. Reşitoğlu, İbrahim Aslan KAK. The pollutant emissions from diesel-engine vehicles and exhaust aftertreatment systems. *Clean Technol Environ Policy*. 2015;17(1):15-27. doi:https://doi.org/10.1007/s10098-014-0793-9.

[13]. Smit R. Development and performance of a new vehicle emissions and fuel consumption software (PΔP) with a high resolution in time and space. *Atmos Pollut Res*. 2013;4(3):336-345. doi:https://doi.org/10.5094/APR.2013.038.

[14]. Cappiello A, Chabini I, Nam EK, Lue A, Zeid MA. A statistical model of vehicle emissions and fuel consumption. In: *Intelligent Transportation Systems, 2002. Proceedings. The IEEE 5th International Conference on*. ; 2002:801-809.

[15]. Cueva-Fernandez G, Espada JP, García-Díaz V, García CG, Garcia-Fernandez N. Vitruvius: An expert system for vehicle sensor tracking and managing application generation. *J Netw Comput Appl*. 2014;42(1):178-188. doi:http://dx.doi.org/10.1016/j.jnca.2014.02.013.

Author's Biography

Bryce Maldonado: Bryce Maldonado is a teaching assistant in the Department of Computer Science and Engineering at the University of South Florida (USA). He is a graduate student with a BS in Computer Science and Engineering having graduated from the College of Engineering at the University of South Florida in 2017.

Malik Bennabi: Malik Bennabi is a senior student of Information Systems Engineering at the University of Sakarya (Turkey). He has been a student of Electrical and Electronic Engineering at the IEEE University of Boumerdès (Algeria). He has completed the Cisco CCNA in 2011. He has been an Erasmus+ scholar at the School of Computer Engineering, University of Oviedo (Spain) during 2016-2017. He has successfully completed several Entrepreneurship programs in different countries. He is a co-founder of the startup I-sen in Asturias (Spain). He is a winner of several entrepreneurship and innovation competitions around the world and has been a mentor at several hackathons including the IoT Olympiads Boston 2014 (USA). He is a polyglot speaking fluently 7 languages, comfortable with 2 and continually learning. He successfully completed the summer research program provided by the University of South Florida (USA) and the University of Oviedo (Spain), summer 2017. His research interests include Machine and Deep learning, Artificial intelligence, Blockchain, Entrepreneurship and Innovative learning.

Vicente García Díaz: Vicente García Díaz is an associate professor in the Department of Computer Science of the University of Oviedo (Spain). He has a PhD from the University of Oviedo in Computer Engineering. He has been a visiting professor in the University of Manchester (UK), Distrital University Francisco José de Caldas (Colombia) and the University of Lisbon (Portugal) and taught courses on computer science in other international institutions. He has published more than 70 research papers in international conferences and journals. He has graduated in prevention of occupational Risks and is a Certified associate in project management through the Project Management Institute. His research interests include domain specific

languages, data science, user of computers for learning and teaching, machine learning and extraction of knowledge from data, extracted from social networks or internet of things-based networks.

Cristian González García: Cristian González García is an assistant professor in the Department of Computer Science of the University of Oviedo (Spain). He is a Technical Engineer in Computer Systems, M.Sc. in Web Engineering, and a Ph.D. in Computers Science graduated from School of Computer Engineering of Oviedo in 2011, 2013, and 2017 (University of Oviedo, Spain). He has been a visitor Ph.D. candidate in the University of Manchester, United Kingdom. Besides, he has been working in different national and regional projects, in projects with private companies, and in the University of Oviedo as a professor. He has published 10 journals articles, 4 conference articles, and 2 book chapters.

Edward Rolando Núñez Valdez: Edward Rolando Núñez Valdez is an assistant professor in the Department of Computer Science of the University of Oviedo (Spain). He is Ph.D. from the University of Oviedo in Computer Engineering. Master in Software Engineering from the Pontifical University of Salamanca and B.S. in Computer Science from Autonomous University of Santo Domingo. He has participated in several research projects, He has taught computer science at various schools and universities and has worked in software development companies and IT Consulting for many years. He has published several articles in international journals and conferences. His research interests include Web Engineering, Artificial Intelligence, Recommendation Systems, Modeling Software with DSL and MDE.

How to Cite

Maldonado, Bryce, Bennabi, Malik, García-Díaz, Vicente, García, Cristian González and Valdez, Edward Rolando Núñez-, "Prediction Model for Pollutants with Onboard Diagnostic Sensors in Vehicles", *International Journal of Machine Learning and Networked Collaborative Engineering,* Vol. 02 No. 2, 2018, pp. 49-57. doi : https://doi.org/10.30991/IJMLNCE.2018v02i02.002.