

Multilayer perceptron method to estimate real-world fuel consumption rate of light duty vehicles

Journal:	IEEE Access
Manuscript ID	Access-2019-03561
Manuscript Type:	Special Section: AI-Driven Big Data Processing: Theory, Methodology, and Applications
Date Submitted by the Author:	30-Jan-2019
Complete List of Authors:	Li, Yawen; Beijing University of Posts and Telecommunications, School of Economics and Management Tang, Guangcan; Hong Kong University of Science and Technology, Department of Computer Science and Engineering Du, Jiameng; Carnegie Mellon University, Electrical and Computer Engineering Wu, Tian; Academy of Mathematics and Systems Science, ; University of the Chinese Academy of Sciences, School of Economics and Management
Keywords:	Artificial intelligence, Data processing, Multilayer perceptrons, Data analysis, Neural networks, Computational and artificial intelligence, Machine learning
Subject Category Please select at least two subject categories that best reflect the scope of your manuscript:	Computational and artificial intelligence, Industry applications
Additional Manuscript Keywords:	big data processing, fuel consumption, light duty vehicle

SCHOLARONE™
Manuscripts

Date of publication xxxx 00, 0000, date of current version xxxx 00, 0000.
Digital Object Identifier 10.1109/ACCESS.2017.Doi Number

Multilayer perceptron method to estimate real-world fuel consumption rate of light duty vehicles

Yawen Li¹, Guangcan Tang², Jiameng Du³, and Tian Wu^{4,5,6}

¹School of Economics and Management, Beijing University of Posts and Telecommunications, Beijing 100876 China
²Department of Computer Science and Engineering, Hong Kong University of Science and Technology, Kowloon 999077 Hong Kong
³Electrical and Computer Engineering, Carnegie Mellon University, Pittsburgh, PA 96801 USA
⁴Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing 100190 China
⁵School of Economics and Management, University of Chinese Academy of Sciences, Beijing 100190 China
⁶Key Laboratory of Big Data Mining and Knowledge Management, Chinese Academy of Sciences, Beijing 100190 China

Corresponding author: Tian Wu (e-mail: wutian@amss.ac.cn).

This work was sponsored by the National Natural Science Foundation of China (71804181). Supported by the National Center for Mathematics and Interdisciplinary Sciences, CAS. Supported by the Fundamental Research Funds for the Central Universities.

ABSTRACT The actual driving condition and fuel consumption rate gaps between lab and real-world are becoming larger. In this paper, we demonstrated an approach to determine the most important factors that may influence the prediction of real-world fuel consumption rate of light duty vehicles. A multilayer perceptron (MLP) method was developed for the prediction of fuel consumption since it provides accurate classification results despite the complicated properties of different types of inputs. The model considers the parameters of external environmental factors, the manipulation of vehicle companies and the drivers' driving habits. Based on the BearOil database in China, 2,424,379 samples were used to optimize our model. We indicated that differences exist between real-world fuel consumption and standard fuel consumption under simulation conditions. This study enables the government and policy-makers to use big data and intelligent systems for energy policy assessment and better governance.

INDEX TERMS Artificial intelligence; big data; multilayer perceptron; fuel consumption rate; light duty vehicles

I. INTRODUCTION

In order to reduce pollution and protect air quality, the promotion of green energy is currently gaining increasing attention in China (Wagner et al., 2009). In recent years, China has issued a series of policies to actively encourage the production and consumption of energy-saving and new energy vehicles. The Ministry of Industry and Information Technology (MIIT), in conjunction with the National Development and Reform Commission, the Ministry of Commerce, the General Administration of Customs and the General Administration of Quality Supervision and Inspection, have drafted the Notice on Strengthening the Management of Average Fuel Consumption in Passenger Vehicle Enterprises in 2014. Furthermore, to achieve the overall goal of the average fuel consumption for passenger vehicles, which is 5 L/100 km in 2020, MIIT has released the revised version of the Limits of Fuel Consumption for Passenger Vehicles and Evaluation Method and Index of

Fuel Consumption of Passenger Vehicles, which states that China will formally implement the fourth phase of fuel consumption standards from January 1, 2016. From 2016 to 2020, the set goal of average fuel consumption for passenger vehicles is 6.7 liters, 6.4 liters, 6 liters, 5.5 liters and 5 liters, respectively, for each year. The standard fully meets the current level of EU standards for automotive oil products, and several indexes even exceed the EU standards. Although great efforts have been made, the automobile industry in China is facing the tremendous pressure of emission reduction (Andre et al., 2006). Generally, the driving cycle fuel consumption data are derived from the website of Automobile Fuel Consumption of China by MIIT, which presents a nearly objective value in the numerical simulation conditions with driving operation templates and reflects the reference samples of real-world fuel consumption data under normal circumstances. It has

been argued that the standard itself has many limitations and vulnerabilities. Because there is a huge difference between the driving conditions in the test procedure for the automobiles and the real world, the automobile manufactures can easily generate very “perfect” data to meet the standard.

There are three main reasons for the difference between standard fuel consumption and actual fuel consumption. First, the external environmental factors, including the heterogeneity of temperature and atmospheric pressure in different regions, and the degree of road traffic congestion (Shi and Abdel-Aty, 2015). As urban road traffic congestion is becoming increasingly serious, car idling may lead to an increase in fuel consumption even if the mileage does not change. Second, another reason for the difference between standard fuel consumption and actual fuel consumption is the manipulation of companies. As automobile companies are familiar with the driving cycle test conditions, they usually provide the vehicles with the best conditions for testing to reduce fuel consumption levels. The third is driving habits, such as air conditioning habits and idle parking. Compared to experienced drivers, new drivers always lack energy-saving abilities during the driving process. Due to the increasing number of vehicle drivers, energy saving is facing more and more severe challenges. As a result, even if a car meets the level of emission standard, it could still be harmful to the environment. Thus, we need to explore the most important factors which may influence the prediction of real-world fuel consumption rate of light duty vehicles.

The rest of the paper is organized as follows: Section II reviews the related literature; Section III presents the methodology and data; Section IV analyzes the results of the multilayer perceptron (MLP) and discusses the policy implications; and finally, the conclusion is presented in Section V.

II. LITERATURE REVIEW

With the rapid development of artificial intelligence (AI), AI-driven big data processing technologies have been used in the field of business activity prediction (Li et al., 2018a; Li et al. 2018b; Zhou et al., 2015). During the development of Plug-in hybrid vehicles (PHEVs), big data and AI models also contributed greatly to estimating the electric grid power demand (Silva et al., 2009). A study helped promote the development of public charging infrastructure by estimating the refueling demand mined from big-data (Cai et al., 2014). Some scholars proposed a range estimation framework through collecting data from the real world to help electric car drivers calculate the remaining driving range (Rahimi-Eichi and Chow, 2014). Furthermore, recent work discusses the challenges of using big data to solve transportation problems, such as the difficulties in collecting and cleaning the data with rich information (Toole et al., 2015). Previous studies also found that intelligent systems can be used in calculating electric vehicle charging demand with the help of

historical traffic data and weather data (Arias and Bae, 2016). Scholars aim to determine the important factors that may predict the innovation efficiency of a NEVs firm. They have built several machine learning models to help firms increase innovation efficiency and build intelligent decision support systems (Li et al., 2018b).

As air-quality protection and energy consumption reduction is gaining attention from emerging economies, an International Vehicle Emissions (IVE) Model was developed to estimate the mobile source emissions in an urban area (Davis et al., 2005). Another IVE model was designed to evaluate the emissions of light duty gasoline trucks, heavy duty gasoline vehicles and motorcycles by utilizing a dataset in China (Guo et al., 2007). To help the government set fuel economy standards for automobiles and reduce energy consumption, scholars have attempted to use the AI-driven (artificial intelligence) big data processing method to create a better test procedure for automobiles, taking the driving situations and environment of the country into consideration. The method can be directly applied to different types of vehicles worldwide (Mahlia et al., 2012).

The government and policy-makers have noticed the important role of big data and intelligent systems for policy assessment and better governance. A previous study built a model to calculate the possibility of an automobile industry in achieving the goal of the European Commission's voluntary agreement under baseline conditions (Zachariadis, 2006). By collecting data from navigation systems, scholars found it possible to analyze the driving and mobility patterns in Europe and calculate the potential of electric vehicles in replacing fuel vehicles. Thus, policies that encourage the development of the next generation of green vehicles could be implemented (Gennaro et al., 2016).

To build next-generation intelligent transportation systems, data processing and mining techniques need to be improved to make full use of real-time information (Zheng et al., 2016). In this study, we present an efficient AI-driven big data processing method to fill this gap, which can better reflect real-world driving conditions. As previous studies indicate that city characteristics, road infrastructure and driving behavior are the most important factors that may influence vehicle emissions, we believe that this study may help government to improve emission estimations and reduce energy consumption in the future (Wang et al., 2008).

III. METHODOLOGY AND DATA

The method we used in this study is multilayer perceptron (MLP). It is a class of feedforward artificial neural networks (ANNs), which are best used for cases when input is high-dimensional and discrete, and the output is real-valued (human readability of the result is unimportant). MLP also reduces the negative impact of possibly noisy data. The main reason we used MLP in this study is that our data might not be linearly separable.

Figure 1 presents the idea that if we have a vector of w values and a vector of x values, and if the dot product of vector w times vector x is positive, then we output a 1; otherwise, we output a 0.

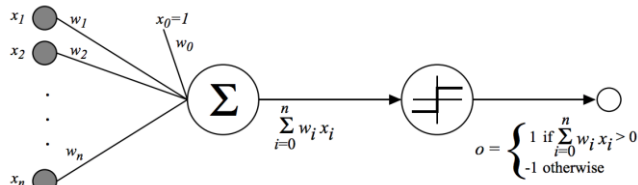


FIGURE 1. Artificial neural network perceptron architecture.

The training rule of perceptron is shown in Equation 1. This rule is guaranteed to succeed if training examples are linearly separable and the learning rate is sufficiently small. The linear unit training rule uses gradient descent, and it is guaranteed to converge to hypothesis with minimum squared error even when training data contain noises.

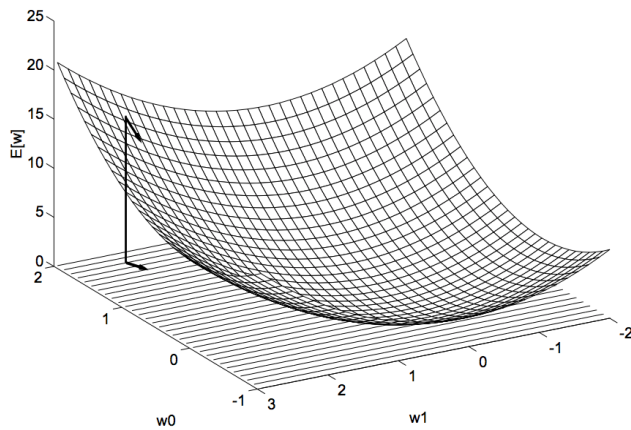


FIGURE 2. Relation between linear unit and squared error.

$$w_i \leftarrow w_i + \Delta w_i \quad (1)$$

where w corresponds to the hidden units, x corresponds to the input features, and each x_i corresponds to an element of the input. $\Delta w_i = \eta(t - o)x_i$, $t = c(\vec{x})$ is the target value, o is the perceptron output, and η is a small constant called learning rate.

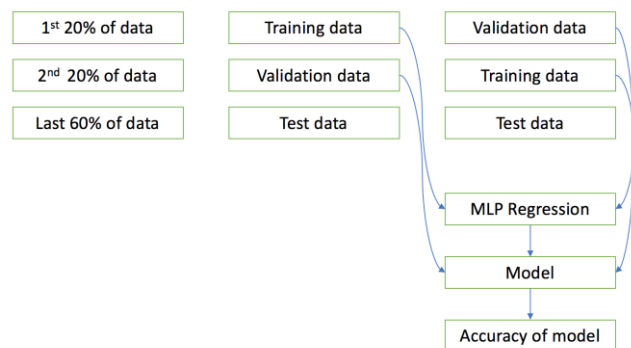


FIGURE 3. Procedure description of prediction.

We used Python to train, validate and test the dataset and imported the 'sklearn' library to split the dataset and train inputs using MLP regression. We then noted the score, average, minimum, maximum and variance of minimum squared errors and the number of hidden layers used.

The real-world fuel consumption data of light duty vehicles used in this paper comes from China's second ranked vehicle and traffic tool (BearOil APP, xiaoxiongyouhao.com). By the end of 2018, the APP had been downloaded 6 million times and has a monthly active user rate of 800 thousand. The accumulated mileage for all active vehicle owners from 31 different autonomous regions in China exceeds 23 billion kilometers, and the number of recorded data for real-world fuel consumption is over 51 million. The factors considered in this paper include host city, vehicle brand, vehicle type, engine parameter, gearbox type and fuel consumption provided by the MIIT. Because the APP also recorded the time each fuel consumption happened, we could use these data to regulate the difference from climate change and accordingly control the variability of the same vehicle based on which city the vehicle owner was in.

We used 2,424,379 samples to optimize our models. Among them, 484,875 samples were used as training data, 484,875 samples were used as validation samples and the rest were considered testing data. We used 20% of the total data as validation to prevent overfitting. If our parameters are fit on the training dataset, then this would result in a biased score. By including another validation dataset, we were able to lock away the test dataset while still being able to measure performance on unseen data as a way of selecting a good hypothesis. In our case, we used the training dataset for learning, fitting the parameters of the classifier and the validating the dataset for considering the number of hidden units in a neural network.

Additionally, we wanted to find the correlation between a certain factor and the result. For a single layer neural network, finding the correlation is simpler. Since the final result is a manipulation of all inputs, we could change each input one at a time and observe how the result changes. For example, if we have input x and input y , and the result is calculated based on $x^2 + y$, then by picking a large y value for the first trial and multiplying by 2 for the second trial, the result would differ by two times as well. However, if we try the same two values on x , then we can observe that the effect on the result is quadratic; therefore, we know the input x should have more impact on the result than y does. This is similar with MLP, only that now we have more inputs, more layers and at each layer, we apply operations to different inputs, so it is harder to observe the relation between certain inputs and the results.

IV. RESULTS

A. DESCRIPTIVE STATISTICAL RESULTS OF REAL-WORLD FUEL CONSUMPTION

Based on the BearOil database, we can see the difference of fuel consumption between MIIT and real-world fuel consumption (See Table I).

TABLE I
DESCRIPTIVE STATISTICAL RESULTS OF FUEL CONSUMPTION

Variable	Obs	Mean	Std.Dev.	Min	Max
MIIT fuel consumption	1,057,046	7.256	1.249	4.200	16.600
Real-world fuel consumption	1,057,046	9.071	2.359	3.000	19.997
Shanghai	216,611	9.487	2.431	3.003	19.997
Lanzhou	10,896	8.839	2.521	3.029	19.924
Beijing	147,335	9.193	2.339	3.000	19.987
Nanning	27,316	8.845	2.310	3.002	19.916
Guangzhou	117,228	8.891	2.352	3.001	19.992
Kunming	29,740	8.065	2.064	3.011	19.869
Hangzhou	61,932	9.252	2.331	3.023	19.992
Wuhan	84,567	9.044	2.234	3.044	19.971
Shenyang	67,017	9.150	2.441	3.006	19.992
Jinan	35,785	8.808	2.427	3.003	19.909
Shijiazhuang	31,907	8.433	2.160	3.052	19.979
Fuzhou	41,931	9.001	2.334	3.000	19.945
Xining	4,952	8.443	2.433	3.002	19.797
Xi'an	53,769	8.804	2.284	3.030	19.994
Guiyang	22,017	9.022	2.248	3.005	19.994
Zhengzhou	44,486	8.856	2.269	3.015	19.966
Changsha	59,557	9.047	2.265	3.011	19.980

From the results, we found that differences exist between real-world fuel consumption and standard fuel consumption released by MIIT. The actual fuel consumption was always higher than the standard consumption. Furthermore, there were obvious regional differences in real-world fuel consumption. The histogram of real-world fuel consumption also shows that the fuel consumption of most vehicle owners was less than 10. Although the standards in China fully meet the current level of the NEDC (New European Driving Cycle) standards for automotive oil products and several indexes even exceed NEDC standards, the NEDC standards may not be treated as the best reference for Chinese energy management. The NEDC standards are always inconsistent with the driving conditions in China. Based on the evaluation of the results from the working condition tests, it has been proven that the actual performance of vehicles in China is quite different from that in European countries. The China Automotive Testing Cycle (CATC) program is being launched by MIIT and other related ministries of China to design the standard of fuel consumption in China.

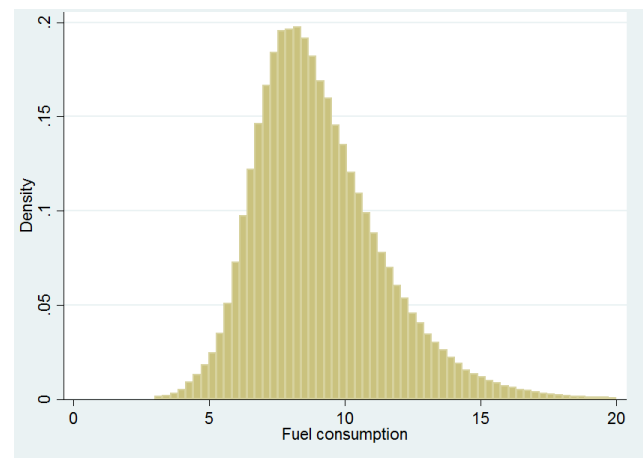


FIGURE 4. Histogram of real-world fuel consumption.

B. RESULTS OF MLP MODEL

Although stated in the methodology section, to evaluate the results of our MLP model, we present the output of our algorithm (See Table II).

TABLE II
RESULT TAKEN FROM MLP MODEL

Hyperparameter	MLP	MLP
Learning Rate	adaptive	adaptive
Hidden Units	20	50
Hidden Layers	20	50
Score	-2.495	-5.404
Average	3,319.215	6,320.949
Maximum	3,498.277	6,600.915
Minimum	3,208.469	6,136.214
Variance	4,460.372	9,972.335

Within the model, we perform 10,000 iterations for the MLP regressor and calculate the minimum squared error for each iteration, and we analyze the performance of our model using the mean accuracy on the given test data and labels. A higher score indicates a better classification of the model; therefore, we pick the model with a lower number of hidden units and hidden layers. We also used the cross-validation method so that we could prevent overfitting the model with our training samples. If our model perfectly fits the training sample but fails to predict the validation sample, then it indicates that our model has a problem of overfitting. With cross validation, we first train using the training sample, then we use the validation sample to check its accuracy in prediction; then, we switch the role of the training sample and validation sample and repeat the above operation. In the end, we average the results. This gives us insight into how our model will generalize to an independent dataset.

To determine which factor influences the final fuel consumption prediction, we could alter the number of factors and see either when the algorithm converges or which results in the greatest decline in classification accuracy. Thus, this

process enables us to obtain a vague idea of which factor contributes most to the final result. However, this approach is not precise because removing any of the inputs will result in a change in the NN architecture and its properties, and the complexity of the sigmoid functions should be considered.

There are other papers about fuel consumption prediction. One used the HDM-III model of the World Bank (Li and Yao, 1992) and analyzed roughness levels of roads and engine characteristics to develop a unit model for fuel consumption. However, that study did not take into account as much data as we did and therefore does not represent as general a case as ours. Another paper used a logistic model to simulate the future trend of China's vehicle population (Wu et al., 2012). However, although this prediction does tell us how much fuel consumption there might be in the future, it could not tell us how much consumption there is for each kind of vehicle with certain characteristics; it just gives us a national trend which is not accurate enough.

Since the MLP method provides accurate classification results despite the complicated property of different types of inputs, we used it for our fuel consumption prediction. Given the nature of MLP, it can be black box that gives little insight to its users of what is actually happening between the layers of hidden units, thus, utilizing sensitivity analysis could raise its transparency and the importance index returned can be used to help us identify which factor affects fuel consumption the most.

V. CONCLUSIONS

In this paper, based on a real-world database that was originally collated for the purpose of fuel consumption prediction, we demonstrated an approach for exploring the main influencing factor on fuel consumption prediction. The model in our study includes the parameters of external environmental factors, the manipulation of vehicle companies and the drivers' driving habits.

Combined with sensitivity analysis, we found that using MLP best classifies the given nonlinear dataset and that the architectures are capable of learning powerful features. To further improve our method, it is easier to split the dataset, other than the test dataset, into more sections to perform a more accurate cross validation. However, such an approach requires our dataset to be larger than it already is, and it requires time for data accumulation.

Additionally, a more precise way of analyzing the importance of each factor for an MLP model is required. As the sensitivity analysis currently only improves the transparency of an MLP model, it is still not able to give us a clear view of how any inputs actually affect the output. In the future, we hope to optimize our model and obtain a better understanding of the relationship between the input factor and output results.

ACKNOWLEDGMENT

This work was sponsored by the National Natural Science Foundation of China (71804181). Supported by the National Center for Mathematics and Interdisciplinary Sciences, CAS. Supported by the Fundamental Research Funds for the Central Universities.

REFERENCES

[1] André, M., Joumard, R., Vidon, R., Tassel, P., & Perret, P. (2006). Real-world european driving cycles, for measuring pollutant emissions from high- and low-powered cars. *Atmospheric Environment*, 40(31), 5944-5953.

[2] Arias, M. B., & Bae, S. (2016). Electric vehicle charging demand forecasting model based on big data technologies. *Applied energy*, 183, 327-339.

[3] Cai, H., Jia, X., Chiu, A. S., Hu, X., & Xu, M. (2014). Siting public electric vehicle charging stations in Beijing using big-data informed travel patterns of the taxi fleet. *Transportation Research Part D: Transport and Environment*, 33, 39-46.

[4] Davis, N., Lents, J., Osses, M., Nikkila, N., & Barth, M. (2005). Part 3: Developing countries: development and application of an international vehicle emissions model. *Transportation Research Record: Journal of the Transportation Research Board*, (1939), 155-165.

[5] De Gennaro, M., Paffumi, E., & Martini, G. (2016). Big data for supporting low-carbon road transport policies in Europe: Applications, challenges and opportunities. *Big Data Research*, 6, 11-25.

[6] Guo, H., Zhang, Q. Y., Shi, Y., & Wang, D. H. (2007). Evaluation of the International Vehicle Emission (IVE) model with on-road remote sensing measurements. *Journal of environmental sciences*, 19(7), 818-826.

[7] Li, X., Yao, Z. (1992). Vehicle fuel consumption prediction model. *Journal of Tongji University*, 20(4), 403-410. (in Chinese)

[8] Li, Y., Jiang, W., Yang, L., & Wu, T. (2018a). On neural networks and learning systems for business computing. *Neurocomputing*, 275, 1150-1159.

[9] Li, Y., Yang, L., Yang, B., Wang, N., Wu, T. (2018b). Application of interpretable machine learning models for the intelligent decision. *Neurocomputing*. In press.

[10] Mahlia, T. M. I., Tohno, S., & Tezuka, T. (2012). A review on fuel economy test procedure for automobiles: Implementation possibilities in Malaysia and lessons for other countries. *Renewable and Sustainable Energy Reviews*, 16(6), 4029-4046.

[11] Rahimi-Eichi, H., & Chow, M. Y. (2014, October). Big-data framework for electric vehicle range estimation. In *IECON 2014-40th Annual Conference of the IEEE Industrial Electronics Society* (pp. 5628-5634). IEEE.

[12] Silva, C., Ross, M., & Farias, T. (2009). Evaluation of energy consumption, emissions and cost of plug-in hybrid vehicles. *Energy Conversion and Management*, 50(7), 1635-1643.

[13] Shi, Q., & Abdel-Aty, M. (2015). Big data applications in real-time traffic operation and safety monitoring and improvement on urban expressways. *Transportation Research Part C: Emerging Technologies*, S0968090X15000777.

[14] Toole, J. L., Colak, S., Sturt, B., Alexander, L. P., Evsukoff, A., & González, M. C. (2015). The path most traveled: travel demand estimation using big data resources. *Transportation Research Part C*, 58, 162-177.

[15] Wagner, D. V., An, F., & Wang, C. (2009). Structure and impacts of fuel economy standards for passenger cars in China. *Energy Policy*, 37(10), 3803-3811.

[16] Wang, Q., Huo, H., He, K., Yao, Z., & Zhang, Q. (2008). Characterization of vehicle driving patterns and development of driving cycles in Chinese cities. *Transportation research part D: transport and environment*, 13(5), 289-297.

[17] Wu, Y., Zhao, P., Zhang, H., Wang, Y., & Mao, G. (2012). Assessment for fuel consumption and exhaust emissions of China's vehicles: Future trends and policy implications. *The Scientific World Journal*, 2012, 1-8.

[18] Zachariadis, T. (2006). On the baseline evolution of automobile fuel economy in Europe. *Energy Policy*, 34(14), 1773-1785.

- [19] Zheng, X., Chen, W., Wang, P., Shen, D., Chen, S., Wang, X., ... & Yang, L. (2016). Big data for social transportation. *IEEE Transactions on Intelligent Transportation Systems*, 17(3), 620-630.
- [20] Zhou, K., Yang, S., Chao, S., Shuai, D., & Sun, C. (2015). Energy conservation and emission reduction of china's electric power industry. *Renewable & Sustainable Energy Reviews*, 45, 10-19.

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

YAWEN LI is an assistant professor of School of Economics and Management, Beijing University of Posts and Telecommunications. She received her Ph.D. in Innovation, Entrepreneurship and Strategy from Tsinghua University in 2018. Her research interest focuses on the collaborative innovation, the development of science parks, and scientific productivity of firms. Her research papers have been published in or accepted by journals including Journal of Cleaner Production, Neurocomputing, Asian Journal of Technology Innovation, Journal of Leadership and Organizational Studies, Chinese Medical Journal.

GUANGCAN TANG is a research assistant in the Department of Computer Science and Engineering, the Hong Kong University of Science and Technology. His research interests include financial technology, machine learning and computer vision.

JIAMENG DU studies in the Electrical and Computer Engineering, Carnegie Mellon University. Her research interests include artificial intelligence, machine learning and data analysis.

TIAN WU is an assistant professor of Academy of Mathematics and Systems Science, Chinese Academy of Sciences. He received his Ph.D. in Finance from Tsinghua University in 2016. His current research topics include operations management, industrial organization, sharing economy, innovation, entrepreneurship and strategy. His research papers have been published in journals including Energy Policy, Energy, Journal of Cleaner Production, Transportation Research Part A: Policy and Practice, Neurocomputing, Resources, Conservation & Recycling, International Journal of Environmental Research and Public Health, Sustainability, Energies.

Fig. 1. Artificial neural network perceptron architecture

Fig. 2. Relation between linear unit and squared error

Fig. 3. Procedure description of prediction

Fig. 4. Histogram of real-world fuel consumption

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

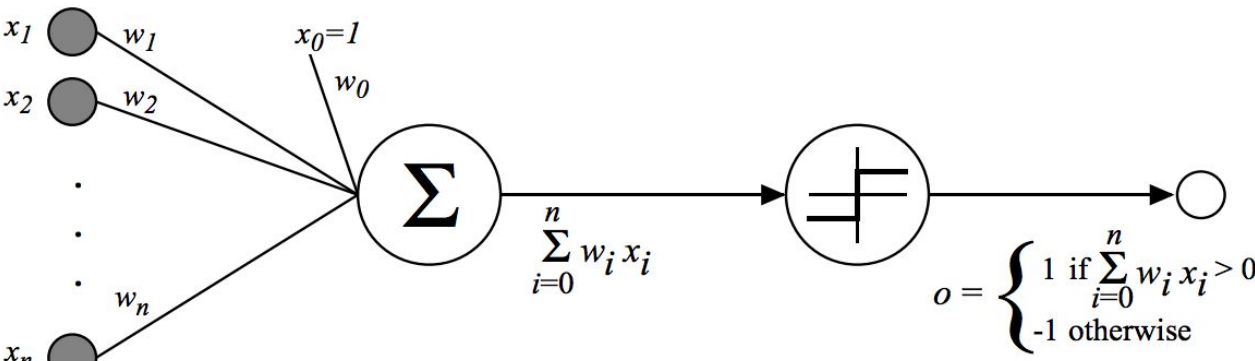


Fig. 1. Artificial neural network perceptron architecture

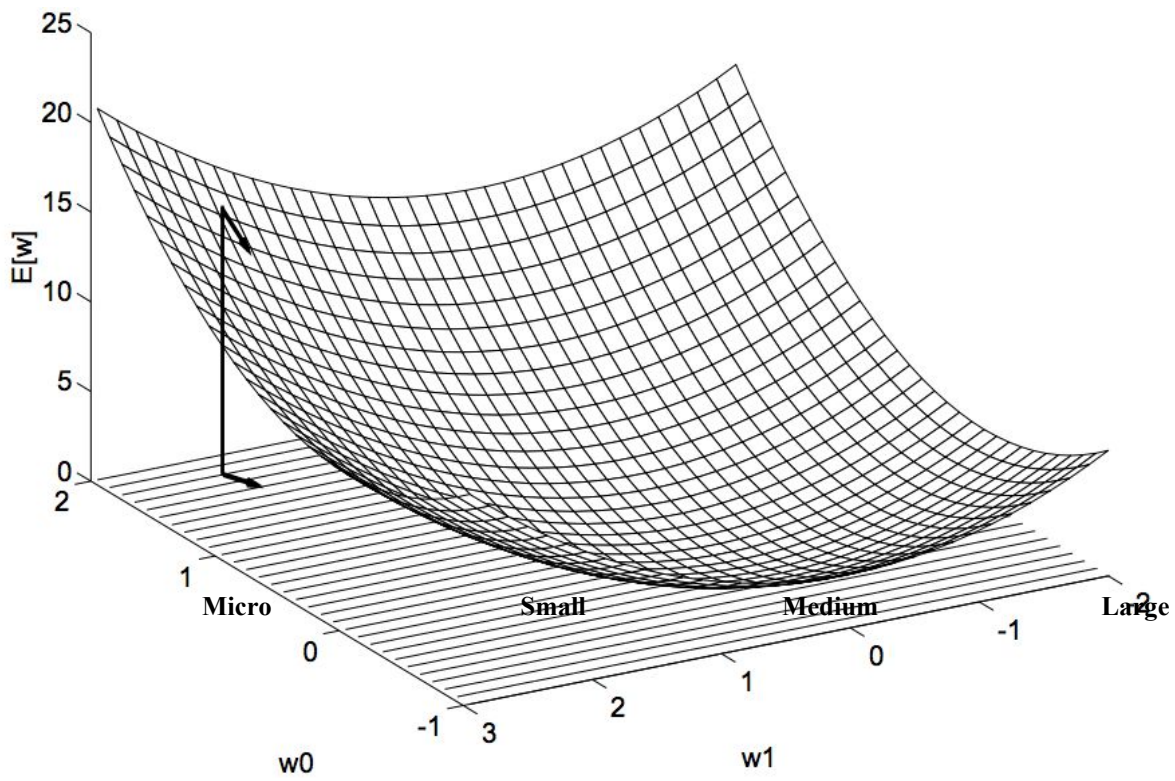


Fig.2. Relation between linear unit and squared error

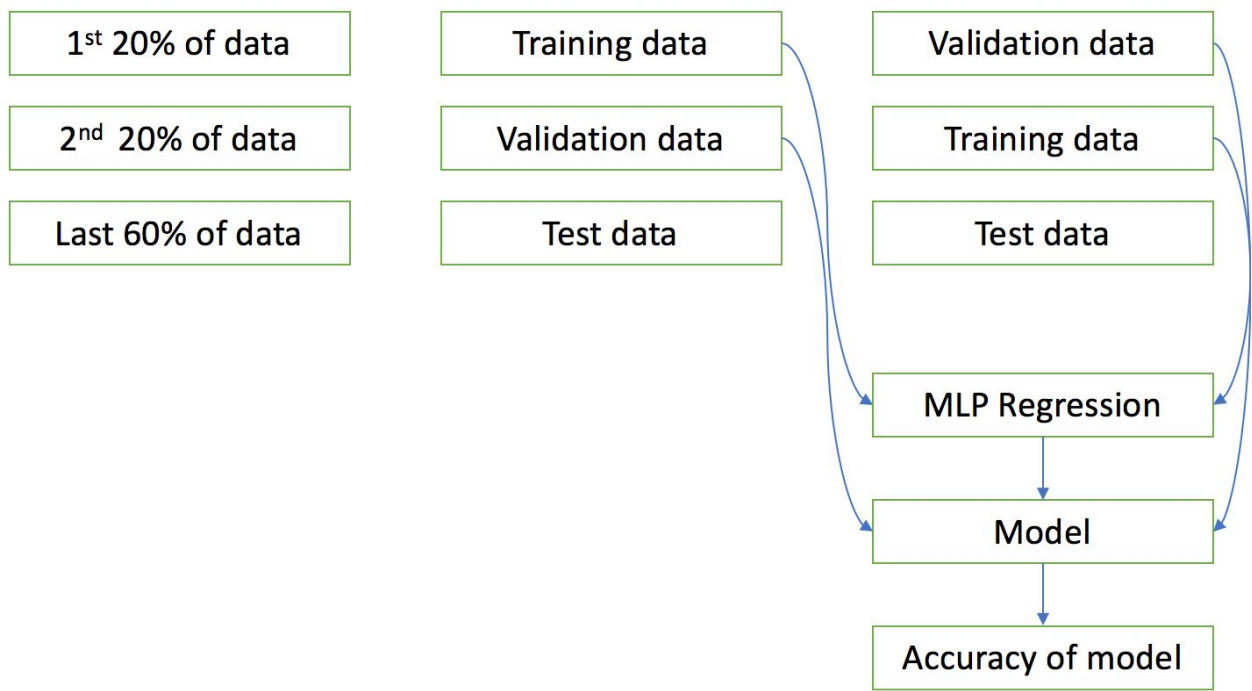


Fig.3. Procedure description of prediction

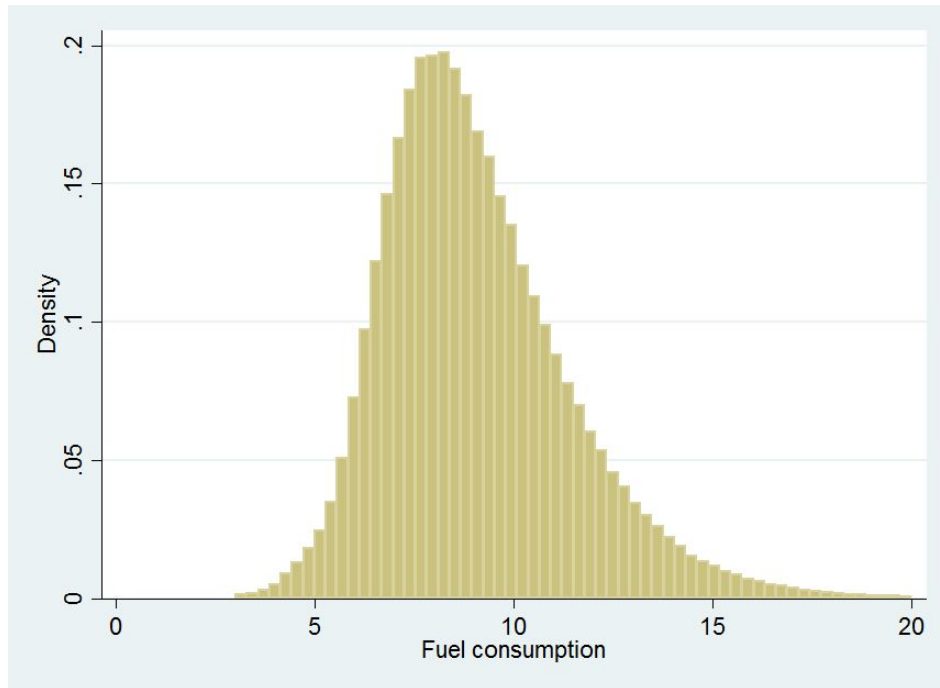


Fig.4. Histogram of real-world fuel consumption

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Table I Descriptive statistical results of fuel consumption.

Table II Result taken from MLP model.

Table I

Descriptive statistical results of fuel consumption.

Variable	Obs	Mean	Std.Dev.	Min	Max
MIIT fuel consumption	1,057,046	7.256	1.249	4.200	16.600
Real-world fuel consumption	1,057,046	9.071	2.359	3.000	19.997
Shanghai	216,611	9.487	2.431	3.003	19.997
Lanzhou	10,896	8.839	2.521	3.029	19.924
Beijing	147,335	9.193	2.339	3.000	19.987
Nanning	27,316	8.845	2.310	3.002	19.916
Guangzhou	117,228	8.891	2.352	3.001	19.992
Kunming	29,740	8.065	2.064	3.011	19.869
Hangzhou	61,932	9.252	2.331	3.023	19.992
Wuhan	84,567	9.044	2.234	3.044	19.971
Shenyang	67,017	9.150	2.441	3.006	19.992
Jinan	35,785	8.808	2.427	3.003	19.909
Shijiazhuang	31,907	8.433	2.160	3.052	19.979
Fuzhou	41,931	9.001	2.334	3.000	19.945
Xining	4,952	8.443	2.433	3.002	19.797
Xi'an	53,769	8.804	2.284	3.030	19.994
Guiyang	22,017	9.022	2.248	3.005	19.994
Zhengzhou	44,486	8.856	2.269	3.015	19.966
Changsha	59,557	9.047	2.265	3.011	19.980

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Table II
Result taken from MLP model.

Hyperparameter	MLP	MLP
Learning Rate	adaptive	adaptive
Hidden Units	20	50
Hidden Layers	20	50
Score	-2.495	-5.404
Average	3,319.215	6,320.949
Maximum	3,498.277	6,600.915
Minimum	3,208.469	6,136.214
Variance	4,460.372	9,972.335