Forming a taxi service order price using neural networks with multi-parameter training

N A Andriyanov¹,²

¹JSC "RPC "Istok" named after Shokin", 2, Vokzalnaya st., Fryazino, Moscow Region, 141190, Russia ²Ulyanovsk State Technical University, 32, Severny Venets st., Ulyanovsk, 432027, Russia

E-mail: nikita-and-nov@mail.ru

Abstract. The article discusses the use of neural networks to forecast the prices of orders in the taxi service. The analysis shows that the results of the price control action are preferable to the results that are obtained for the direct cost of the trip. Special attention is paid to such factors as the time of the order, the weather condition of the environment, the number of cars in the area, as well as the distance between the starting and ending points of the route. It was shown that networks ensure maximum accuracy with backward propagation of error with the number of neurons (5) close to the number of parameters (4).

1. Introduction

Artificial intelligence can now be successfully applied in various fields of human activity [1–8]. In [9–11], the problem of forecasting the number of taxi service orders was considered. However, in addition to forecasting the number of orders, tariffing also plays an important role. Typically, taxi service billing systems work according to a simple algorithm, often without taking into account the many possible factors for pricing a trip. The paper discusses the use of a trained neural network to form the price of a trip, taking into account many factors.

2. Training sample collecting

For the study the pricing algorithm based on neural networks, the following factors were selected as significant factors: order time, probability of rainfall, minimum trip cost, number of free cars in the order area, trip distance. Accordingly, a control action was formed at the output that affected only the minimum cost of the trip in these conditions, as well as the price of the trip itself. Decisions were made by taxi order managers.

Figures 1-4 show the dependence of the order price distribution separately, depending on each parameter. Figure 1 describes the relationship between price and order time (normalized from 0 to 1), Figure 2 describes the relationship between the price and the effect of precipitation (normalized from -1 to 1), Figure 3 describes the relationship between the price and the number of nearby cars (normalized from 0 to 1) Figure 4 describes the relationship between price and distance (not normalized, in km). The price always lays along the Y-axis and normalized to some average value.
Figure 1. Price distribution by order time

Figure 2. Price distribution by precipitation level

Figure 3. Price distribution by the number of waiting cars in the area
The presented dependencies allow concluding that the most important factor affecting the formation of prices is the distance of the trip. However, a more informative picture of the connections can be obtained from Figures 5-8, where everything is similar to Figures 1-4, but for a control action on the price.
Figure 6. Control action distribution by precipitation level

Figure 7. Control action distribution by the number of waiting cars in the area
Figure 8. Control action distribution by distance

In this form, the possibility of some general connection of the control action with each of the parameters is really assumed.

Thus, it is necessary to train the network both for forecasting the control action and for forecasting the total cost of the order (price).

3. Experiment results

For train the neural network, 122 vectors were created with 5 input parameters and 2 output parameters. The training sample was 86 vectors of each parameter for validation and 18 vectors of each parameter for the test. Networks with back propagation of error were used, as well as Elman networks and general regression networks.

The results of the relative mean square error are presented in Table 1 for forecasts of the control action and price for the training sample and test for different numbers of neurons in the network.

| Network                | MSE(training dataset) | MSE (test dataset) |
|------------------------|------------------------|--------------------|
| FFBP1 for control action| 0.1229                 | 0.1325             |
| FFBP 1 for price       | 0.2084                 | 0.1535             |
| FFBP3 for control action| 0.0480                 | 0.0277             |
| FFBP3 for price        | 0.0292                 | 0.0577             |
| FFBP5 for control action| **0.0249**            | **0.0176**         |
| FFBP5 for price        | 0.1302                 | 0.0838             |
| FFBP10 for control action| 0.0955                | 0.0341             |
| FFBP 10 for price      | 0.0427                 | 0.0658             |
| FFBP50 for control action| 0.1618                | 0.3290             |
| FFBP50 for price       | 0.3302                 | 0.5729             |
| ENN10 for control action| 0.1091                | 0.1055             |
| ENN 10 for price       | **0.0225**            | **0.0501**         |
| GR for control action  | 0                      | 0.5844             |
| GR for price           | 0                      | 0.2579             |
Here FFBP (Feed Forward Back Propagation) is the network with the back propagation of the error, GR (General Regression) is the network of general regression. ENN is the Elman neural network. The number of neurons is indicated after the network name. As can be seen from Table 1, a network based on 5 neurons best approximates data on the formation of control action. For the formation of prices, the Elman network of 10 neurons works best. Even though general regression networks provide absolute accuracy in the training set, their application in the test set leads to significant errors.

Figure 9 shows the forecast of the control action of the test sample: the solid line is the real value, the dashed line is the worst network, and the crosses are the best network.

4. Conclusion
The use of neural networks for pricing in a taxi order service is proposed depending on various conditions. It is shown that the results obtained with a small number of neurons allow a fairly accurate approximation of the wishes of managers. At the same time, such pricing may allow individual customers saving on travel and paying more. However, for the orders submitted, the total cost increased by 2% compared with the linear pricing algorithm, which is an acceptable deviation.

5. Acknowledgments
This work was supported by the RFBR Grants, Projects No. 19-29-09048 and No. 19-47-730011

References
[1] Dadashev Z.F., Ustinova H.G. 2019 The Age of Science 18 12-24
[2] Ivanov A.A., Rozhkova L. 2018 Bulletin of St. Petersburg State University of Economics 3(111) 24-32
[3] Schurina S.V., Danilov A.S. 2019 Economics. Taxes. Right.3 139-145 2019
[4] Sinitso S.A. 2019Bulletin of Eurasian science 2 6-14
[5] Zoidov K. K., Ponomareva S.V., Serebryansky D.I. 2018 RPPE 9(95) 36-45
[6] Teterin D.A., Khabibulin R.Sh., Gudin S.V. 2018 Scientific reports of Belgorod State University. Series: Economics. Informatics 3 8-12
[7] Begishev I.R., Khisamova Z.I. 2018 All-Russian Criminological Journal 6 21-27
[8] Sukhodolov A.P., Bychkova A. 2018 All-Russian Criminological Journal 6 32-40
[9] Danilov A.N., Andriyanov N.A., Azanov, P.T. 2018 Journal of Physics: Conference Series 1096 1-8 DOI:https://doi.org/10.1088/1742-6596/1096/1/012188
[10] Azanov P., Danilov A., Andriyanov N. 2017 CEUR Workshop Proceedings 1904 232-238 DOI:10.18287/1613-0073-2017-1904-232-238
[11] Andriyanov N.A., Sonin V.A. 2018 CEUR Workshop Proceedings 2258 462–472