https://doi.org/10.37528/FTTE/9788673954165/POSTEL.2022.008

TESTING FOR IMPROVEMENT IN PREDICTION MODEL PERFORMANCE IN POST

Nataša Milosavljević[1], Nikola Knežević[2], Nebojša Bojović[2], Miloš Milenković[2]
[1]University of Belgrade - Faculty of Agriculture,
natasam@agrif.bg.ac.rs
[2] University of Belgrade - Faculty of Transport and Traffic Engineering,
n.knezevic@sf.bg.ac.rs, nb.bojovic@sf.bg.ac.rs, m.milenkovic@sf.bg.ac.rs

Abstract: *The aim of this research is to compare the performance of predictive models when different metaheuristics are used. We improved the neural networks using the bee algorithm (BCO) and the ant colony algorithm (ACO) on the shipment data. They compared the obtained results and found that the model of the BCO algorithm performed better than the ACO, and both better than the basic model of neural networks. In further research, our goal is to achieve even better performance by combining several algorithms in one. For the model of neural networks without using metaheuristics, we used R, and we implemented the other models in the programming language python.*

Key words: *Prediction, machine learning, hybrid models, meta-heuristic algorithms, artificial intelligence.*

1. Introduction

Intelligent Transportation System (ITS) aims to increase the operational efficiency and capacity of the transportation system by creating an integrated system of people, roads and vehicles (An et al., 2011). An efficient ITS environment requires a continuous flow of information regarding how traffic conditions evolve with time (Lieu, 2000), and one of the most frequently studied conditions is traffic flow (Vlahogianni et al., 2004), which refers to the number of vehicles passing through a given point on a road segment in a certain time span. Traffic flow forecasting can be used for specific tasks ranging from road condition controlling (Jiang and Adeli, 2005) to travel planning (Lee et al., 2009), hence is strongly needed for individual road users, business sectors, and government agencies.

A predominant change in ITS recently is extensive data can be collected from various sources (Zhang et al., 2011), which prompts the prevalence of data-driven methods for traffic forecasting. Unlike knowledge-driven methods employing analytical or simulation models (Cascetta, 2013), data-driven approaches develop models directly learning the traffic dynamics from traffic data, and are generally more accurate and robust (Van Lint and Van Hinsbergen, 2012). From the perspective of the forecasting period, data-driven methods can be classified into short-term (from a few seconds to a few hours) (Vlahogianni et al., 2014) and long-term (longer than short-term and up to 24 hours) forecasting (Hou et al., 2015), and the former has attracted most effort till several years ago (Vlahogianni et al., 2014).

The goal of this research is to deal with predictive models based on the data of the Post of Serbia. We will not follow the problems of transport itself, which is connected with letters, packages and express shipments, which will be the goal of our research. The work is based on the description of the methods that we used in this research, and then we will describe the data, present the results and give the conclusion of this research.

2. Methodology

Traditionally, NNs models learn by changing the interconnection weights of their associated neurons. It can be trained by different approaches such as: BP, Improved BP algorithm, Evolutionary Algorithms (EA), Swarm Intelligence (SI), Differential Evolution (DE) and Hybrid Bee Ant Colony (HBAC), IABC-MLP and recently HABC Algorithms. However, BP algorithm results in long training time and insufficient performance for the binary classification task. However, a BP learning algorithm has some difficulties; especially, it's getting trapped in local minima, where it can affect the NNs performance (Bonabeau et al. 1999.).

NN learning is a process of obtaining new knowledge or adjusting the existing knowledge through the training process. The combination of weights, which minimizes the error function is considered to be a solution of the learning problem. This step by step mathematical procedure adjusts the weights according to the error function. So, the adjustment of weights, which decrease the error function is considered to be the optimal solution of the problem. In the input layer only inputs propagate through weights and passing through hidden layers and get output by some local information. For the BP error, each hidden unit is responsible for some part of the error.

Ant Colony Optimization (ACO) is a meta-heuristic procedure for the solution of a combinatorial optimization and discrete problems that has been inspired by the social insect's foraging behaviour of real ant decision developed in 1990s. Real ants are capable of finding Food Source (FS) by a short way through exploiting pheromone information, because ants leave pheromone on the ground, and have a probabilistic preference for trajectory with larger quantity of pheromone. Ants appear at a critical point in which they have to choose to get food, whether to turn right or left. Initially, they have no information about which is the best way for getting the FS.

Ants move from the nest to the FS blindly for discovering the shortest path. The above behavior of real ants has inspired ACO, an algorithm in which a set of artificial ants cooperate in the solution of a problem by sharing information. When searching for food, ants initially explore the area surrounding their nest in a random manner. As soon as an ant finds FS, it evaluates the quantity and the quality of the food and carries some of it back to the nest. The following is the ACO pseudo code.

```
Initialize Trail
Do While (Stopping Criteria Not Satisfied) – Cycle Loop
Do Until (Each Ant Completes a Tour) – Tour Loop
Local Trail Update
End Do
Analyze Tours
Global Trail Update
End Do
```

Artificial Bee Colony (ABC) Algorithm was proposed for optimization, classification, and NNs problem solution based on the intelligent foraging behavior of honey bee. The three bees determine the objects of problems by sharing information to other's bees. The employed bees use multidirectional search space for FS with initialization of the area. They get news and all possibilities to find FS and solution space. Sharing of information with onlooker bees is performed by employed bees. Onlooker bees: Onlooker bees evaluate the nectar quantity obtained by employed bees and choose a FS depending on the probability values calculated using the fitness values. Onlooker bees watch the dance of hive bees and select the best FS according to the probability proportional to the quality of that FS. Scout bees: Scout bees select the FS randomly without experience. If the nectar quantity of a FS is higher than that of the old source in their memory, they memories the new position and forget the previous position. Whenever employed bees get a FS and use the FS very well again, they become scout bees to find a new FS by memorizing the best path.

Bee colony and ant colony we used to get an arrangement of neurons in the middle layer that would give better results than a regular neural network algorithm.

3. Results and discussion

In our research, we used data from the Post of Serbia. We had three time series: letters, packages, and express packages. Letters (Figure 1.) and packages (Figure 2.) had 83 observations, covering the period from 2001 to 2022 (4 data for each year, with 2001 having 1 and 2022 having 2 data). Express packages (Figure 3) had 57 observations and covered the period from 2008 to 2022 (3 data in 2008 and 2 in 2022).

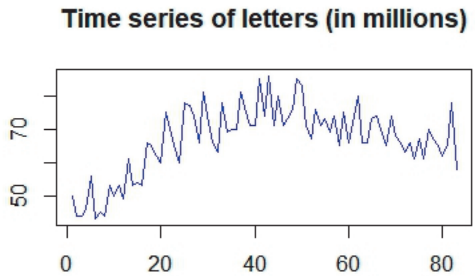

Figure 1. Time series observation for letters from 2001. to 2022.

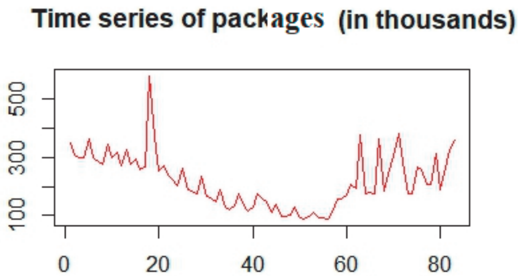

Figure 2. Time series observation for packages from 2001. to 2022.

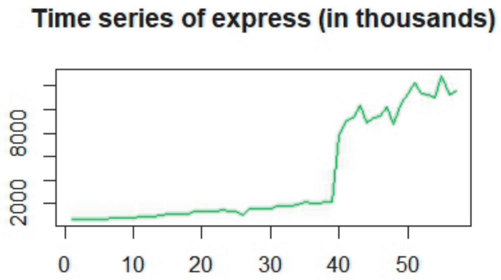

Figure 3. Time series observation for express packages from 2008. to 2022.

The neural network used in the first two time series 75 observations were for training the neural network, and 8 for prediction validation. In the third, we used 50 for training the neural network, and 7 for validating the prediction.

We varied the number of neurons in the middle layer. It went from 5 to 15. The criterion, both with the simple and with the combination of models, was the smallest prediction error.

When looking at a time series of letters the smallest prediction error was not achieved in the model to be the same in each model for the same number of neurons in the middle layer. In the case of the ordinary neural network, this was achieved with 7 neurons in the middle and it was 0.31, in the dark colony it was with 9 neurons in the middle and it was 0.22, and in the case of the bee colony 5 neurons in the middle and it was 0.20. The criterion we decided to show the results and compare these three models was to select the number of neurons in the middle where all three models had the best prediction result (in terms of average model success value).

The bee colony model generally performed best, then the ant model and finally the regular neural network model. According to the criteria we decided on, the models showed the best prediction for 8 neurons in the middle layer (Figure 4.).

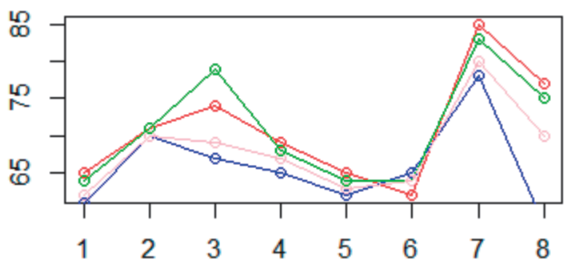

Figure 4. The model for letters that proved to be the best. Blue represents the original data, red the neural network, green the ant colony, and pink the bee colony.

The bee model also performed best with the package. There, the best result was shown with 10 neurons in the middle layer (Figure 5.), while in particular, the best result was achieved by a neural network with 8 neurons in the middle, an ant colony with 12, and bees with 10 neurons in the middle. In this case, the model errors were: neural networks 0.34, ant colony 0.29, and bee algorithm 0.27.

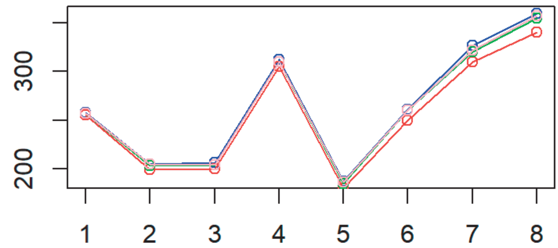

Figure 5. The model for packages that proved to be the best. Blue represents the original data, red the neural network, green the ant colony, and pink the bee colony.

The situation is similar with regard to the model of the express package. The best was the bee algorithm with an error of 0.12, followed by the ant algorithm with 0.26 and finally the neuron algorithm with 0.41 (Figure 6.). The number of neurons in the middle layer for this case was 7. If we look at the individual behavior of the error movement here as well, the neural networks were the most successful for 12 neurons in the middle, the ant colony for 9, and the bee algorithm for 10.

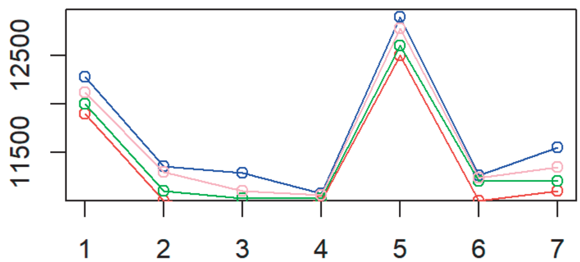

Figure 6. The model for express packages that proved to be the best. Blue represents the original data, red the neural network, green the ant colony, and pink the bee colony.

4. Conclusion

From the results achieved in this research, as well as from the results of other researchers who dealt with similar problems, combining different models proved to be better than traditional models.

In our research, the hybrid model that used bees to improve the standard neural network proved to be better, but in general it is necessary to examine the variation of not only the neurons in the middle layer, but also to use the advantages of metaheuristic algorithms and to vary their parameters together with the parameters of the neural network in order to achieve even better results. Our future research will go exactly in this direction.

Acknowledgement

This work was created as a result of research within the Agreement on the implementation and financing of scientific research work in 2022 between the Faculty of Agriculture in Belgrade and the Ministry of Education, Science and Technological Development of the Republic of Serbia, contract registration number: 451-03-68/2022-14/200116.

Literature

[1] AN, S.-H., LEE, B.-H. & SHIN, D.-R. A survey of intelligent transportation systems. 2011.
[2] LIEU, H. C. 2000. Traffic estimation and prediction system.
[3] VLAHOGIANNI, E. I., GOLIAS, J. C. & KARLAFTIS, M. G. 2004. Short□term traffic forecasting: Overview of objectives and methods. Transport reviews, 24, 533-557.
[4] JIANG, X. & ADELI, H. 2005. Dynamic wavelet neural network model for traffic flow forecasting. Journal of transportation engineering, 131, 771-779.
[5] LEE, W.-H., TSENG, S.-S. & TSAI, S.-H. 2009. A knowledge based real-time travel time prediction system for urban network. Expert systems with Applications, 36, 4239-4247.
[6] ZHANG, J., WANG, F.-Y., WANG, K., LIN, W.-H., XU, X. & CHEN, C. 2011. Data-driven intelligent transportation systems: A survey. IEEE Transactions on Intelligent Transportation Systems, 12, 1624-1639.
[7] CASCETTA, E. 2013. Transportation systems engineering: theory and methods, Springer Science & Business Media.
[8] VAN LINT, J. & VAN HINSBERGEN, C. 2012. Short-term traffic and travel time prediction models. Artificial Intelligence Applications to Critical Transportation Issues, 22, 22-41
[9] VLAHOGIANNI, E. I., KARLAFTIS, M. G. & GOLIAS, J. C. 2014. Short-term traffic forecasting: Where we are and where we're going. Transportation Research Part C: Emerging Technologies, 43, 3-19.
[10] HOU, Y., EDARA, P. & SUN, C. 2015. Traffic Flow Forecasting for Urban Work Zones. IEEE Transactions on Intelligent Transportation Systems, 16, 1761-1770.
[11] BONABEAU, E., DORIGO, M., THERAULAZ, G.: Swarm Intelligence: From Natural to Artificial Systems. Oxford University Press, NY (1999).

Sadržaj: *Cilj ovog istraživanja je da uporedi performanse prediktivnih modela kada se koriste različite metaheuristike. Neuronske mreže smo unapredili korišćenjem pčelinjeg algoritma i algoritma kolonije mrava na podacima pošiljki. Uporedili dobijene rezultate i dobili da se bolje pokazao model pčelinjeg algoritma od kolonije mrava, a oba bolje od osnovnog modela neuronskih mreža. U daljim istraživanjima cilj nam je da kombinacijama više algoritama u jednom postignemo još bolje performanse. Za model neurnoskih mreža bez korišćenja metaheuristike koristili smo R, a ostale modele smo implementirali u programskom jeziku python.*

Ključne reči: *Predikcija, mašinsko učenje, hibridni modeli, metaheuristički algoritmi, veštačka inteligencija.*

TESTIRANJE MODELA PREDVIĐANJA U CILJU POBOLJŠANJA PERFORMANSI U POŠTI

Nataša Milosavljević, Nikola Knežević, Nebojša Bojović, Miloš Milenković