Artificial Neural Networks for forecasting passenger flows on metro lines

Mariano Gallo, Giuseppina De Luca, Luca D’Acierno, and Marilisa Botte

Abstract—Forecasting users’ flows on transportation networks is a fundamental task of Intelligent Transport Systems (ITSs); indeed, most control and management strategies on transportation systems are based on the knowledge of user flows. For implementing ITS strategies, the forecast of user flows on some network links obtained as a function of user flows on other links (for instance, where the data are available in real-time) may be a significant contribution. In this paper, we propose the use of Artificial Neural Networks (ANNs) for forecasting metro onboard passenger flows as a function of passenger counts at stations’ turnstiles. We assume that metro station turnstiles allow collecting the number of entering passenger by means of an automatic counting system and these data are available every some minutes (temporal aggregation); the objective is to estimate on-board passengers on each track section of the line (i.e. between two successive stations) as a function of turnstiles’ data collected in the previous periods. The choice of the period length may depend on service schedules. Artificial Neural Networks are trained by using simulation data obtained with a dynamic loading procedure of the rail line; the proposed approach is tested on a real-scale case: the Line 1 of the Naples metro system (Italy). Numerical results show that the proposed approach is able to forecast the flows on metro sections with satisfactory precision.

Keywords—Artificial Neural Networks, metro, transportation, user flow forecast.

I. INTRODUCTION

The knowledge of user flows on transportation systems is crucial for implementing control and management policies. In this context, the monitoring systems assume a central role and are diffused in many road and rail networks: they are one of the most important (and necessary) components of Intelligent Transport Systems (ITSs).

Monitoring systems are based on sensors (or detectors) that measure some characteristics of flows and transmit them to a control room, where these data are used for implementing control and management strategies. For instance, to know traffic flows on a road network is useful for implementing flow-responsive traffic-signal systems, while the user loads on public transport vehicles can be used for real-time scheduling/rescheduling tasks.

Despite their usefulness, monitoring systems are not always provided on road and rail networks or, sometimes, implemented systems do not have enough sensors to collect the data necessary to control and management strategies. Indeed, the costs of these systems require significant investments from public administrations or public transport firms that, often, are not admissible.

In this paper, we focus on metro lines, where the accesses are controlled with turnstiles, which can count the passengers entering each station, with or without identifying their direction (in most of the existing stations, except the terminals, passengers can board on trains with different directions). These data can be easily collected every some minutes (e.g., fifteen-minute intervals) without installing new sensors. The objective of this paper is to propose a method for estimating the number of passengers on each segment of a metro line, knowing the data at turnstiles.

In the following, we propose to use Artificial Neural Networks (ANNs) for estimating the onboard passengers of a metro line, using as input data the counts at turnstiles.

The paper is organised as follows: section II examines the background; section III describes the problem to solve and the ANN approach; the method adopted for generating the data used in numerical tests is reported in section IV; section V describes the case study and the results; section VI concludes...
and identifies some research prospects.

II. BACKGROUND

A. Artificial Neural Networks

The Artificial Neural Network (ANN) is a mathematical method that is widely used for reproducing several physical phenomena and forecasting the results of some actions on (or variations of) the parameters/variables of the system. ANNs are considered black-boxes since the functions and the relationships between inputs and outputs are hidden, not known and, generally, not interpretable.

Both the strengths and weaknesses of ANNs are related to their black-box approach. ANNs can reproduce a phenomenon or approximate a function without making explicit the parameters; moreover, once trained, they are able to give the results rapidly. On the other hand, the trained ANNs are not extendible even to similar cases and work only if the boundary conditions do not change significantly.

ANNs were introduced in [1]-[3]; other pioneering contributions can be found in [4]-[8]. Some general books are [9]-[14], while some literature reviews can be found in [15]-[21].

B. Transportation flow forecasting and ANN approaches

Two main types of transportation flow forecasting problems can be identified: i) short-term forecasting and ii) traffic data spatial extension. A review can be found in [22].

The first problem aims to forecast the traffic flows (or user flows) that cross a road section (or use a transit line) in a future time interval, using the data measured in the previous time intervals in the same section. This problem is widely studied in the literature and a complete review should deserve a specific paper; hence, we refer to [23], [24] that have reviewed the problem. This problem has been tackled with several approaches, and the ANNs have been widely used too [25]-[44].

The second problem, instead, aims to estimate the traffic flows (or user flows) on some components of a network (e.g., roads links, transit links, etc.), using the data measured on other components. This problem was less studied in the literature [43], [44]; recently, Gallo and De Luca [45] proposed an ANN approach that has given good results on a real-scale road network.

III. PROBLEM DESCRIPTION AND ANN APPROACH

We assume that turnstiles control all accesses to a metro line: each user, entering a station, uses a ticket (or a pass) for crossing the turnstile. Moreover, the turnstiles are able only to count users entering the station without linking the origin of each trip with the corresponding destination. This situation is common in many metro lines, as Line 1 of the Naples metro system (Italy) that will be the object of the real-scale test. Often, indeed, the turnstiles are installed only for facilitating the ticket control/validation and avoiding no-ticket trips, and, in urban contexts, the ticket is the same regardless of the origin-destination pair. In the following, we will consider two cases: a) turnstiles at the station entrance that measure only the entering passengers, without any indication on the direction that they will follow; b) turnstiles at the accesses to the platforms that give information also on the trip direction.

The data collected by turnstiles can be used, with a low technologic investment, for implementing a monitoring system of the whole metro line, generating information about the passengers on each railway section (between two stations). This information can be very useful to metro operators for implementing real-time strategies, such as a frequency increase or reduction, the definition of train composition (number of passenger carriages), the scheduling of additional runs, and so on.

The problem to solve is the estimation of loads on the line starting from turnstile data. For solving this problem, we propose to use feedforward ANNs, which are suitable because a) it is not necessary to explicitly know the relationships between inputs and outputs, b) the results are obtained rapidly, and c) usually, the boundary conditions do not change so much to invalidate the forecasts.

The structure of the ANN provides an input layer with a node for each turnstile datum, an output layer with a node for each convoy load datum, and one or more hidden layers. The best structure of the ANN has to be designed for each specific problem. A crucial point is related to the dynamic nature of the problem: indeed, the train moves along the line and loads and unloads passengers at stations in different time intervals. Therefore, onboard passengers between two stations at time $t$ depends on the passengers loaded and unloaded to the previous stations at different times $\lt t$. For this reason, the inputs of the ANN have to regard turnstile counts referring to more time intervals preceding the ones under forecast. The number of inputs will depend on the travel time duration between terminals.

The other crucial point is the training phase of ANNs; here we propose to use a supervised learning method, where the example datasets are generated through dynamic simulation models (see section IV).

IV. GENERATION OF TRAINING DATASETS

To generate the training datasets, we use the simulation model proposed in [46]; this model assumes that:

1) platforms can accommodate all passengers (incoming, waiting and outgoing);
2) at each station and for each direction there is only one platform available;
3) the dwell time is constant and independent on the number of passengers alighting and boarding;
there is no interaction between alighting, boarding and waiting passengers;

5) the capacity of each train is fixed;

6) passengers are distributed uniformly among the train coaches;

7) there is no interaction in the train between alighting, boarding and onboard passengers.

In our test, we assume that the passengers follow a FIFO (First In-First Out) rule for boarding the convoy. The analytic details of the model can be found in [46].

Using this model, we generate the training datasets on the case study as follows: a) numerous origin-destination (OD) matrices referring to 15’ intervals are randomly generated starting from a base OD matrix; b) four OD matrices, referring to four time intervals, are assigned to the metro line, giving as results the passengers counted at turnstiles, for each interval, and the passengers onboard in each railway section in the next time interval; c) the output data of the problem are the passengers on railway sections, while the input data are the flows on railway sections between 10:15 and 10:30 are calculated in function of the turnstile counts in the following intervals: 9:15-9:30, 9:30-9:45, 9:45-10:00 and 10:00-10:15). Therefore, the structure of the training datasets is reported in Fig. 1, where the following notations are used:

\( ds \) is the number of datasets;

\( t \) is the period under analysis;

\( c^t_{j,i} \) is the passenger count at turnstile \( j \) in period \( t \) for dataset \( i \);

\( f^t_{k,i} \) is the load on railway section \( k \) in period \( t \) for dataset \( i \).

### Table 1

| Dataset \( \rightarrow \) | 1 | 2 | ... | \( ds \) |
|---------------------------|---|---|-----|-----|
| Period \( t = \) | \( t_{in,ds} \) | \( t_{out,ds} \) | ... | ... |

**Input data**

- Turnstile 1 - Period t-1: \( c^t_{1,j,1} \), \( c^t_{1,j,2} \), ... , \( c^t_{1,j,d} \)
- Turnstile 2 - Period t-1: \( c^t_{2,j,1} \), \( c^t_{2,j,2} \), ... , \( c^t_{2,j,d} \)
- ... for each turnstile
- Turnstile ts - Period t-1: \( c^t_{ts,j,1} \), \( c^t_{ts,j,2} \), ... , \( c^t_{ts,j,d} \)
- Turnstile 1 - Period t-2: \( c^{t-1}_{1,j,1} \), \( c^{t-1}_{1,j,2} \), ... , \( c^{t-1}_{1,j,d} \)
- Turnstile 2 - Period t-2: \( c^{t-1}_{2,j,1} \), \( c^{t-1}_{2,j,2} \), ... , \( c^{t-1}_{2,j,d} \)
- ... for each turnstile
- Turnstile ts - Period t-2: \( c^{t-1}_{ts,j,1} \), \( c^{t-1}_{ts,j,2} \), ... , \( c^{t-1}_{ts,j,d} \)
- Turnstile 1 - Period t-3: \( c^{t-2}_{1,j,1} \), \( c^{t-2}_{1,j,2} \), ... , \( c^{t-2}_{1,j,d} \)
- Turnstile 2 - Period t-3: \( c^{t-2}_{2,j,1} \), \( c^{t-2}_{2,j,2} \), ... , \( c^{t-2}_{2,j,d} \)
- ... for each turnstile
- Turnstile ts - Period t-3: \( c^{t-2}_{ts,j,1} \), \( c^{t-2}_{ts,j,2} \), ... , \( c^{t-2}_{ts,j,d} \)
- Turnstile 1 - Period t-4: \( c^{t-3}_{1,j,1} \), \( c^{t-3}_{1,j,2} \), ... , \( c^{t-3}_{1,j,d} \)
- Turnstile 2 - Period t-4: \( c^{t-3}_{2,j,1} \), \( c^{t-3}_{2,j,2} \), ... , \( c^{t-3}_{2,j,d} \)
- ... for each turnstile
- Turnstile ts - Period t-4: \( c^{t-3}_{ts,j,1} \), \( c^{t-3}_{ts,j,2} \), ... , \( c^{t-3}_{ts,j,d} \)

**Output data**

- Railway section 1 - Period t: \( f^t_{1,k,1} \), \( f^t_{1,k,2} \), ... , \( f^t_{1,k,d} \)
- Railway section 2 - Period t: \( f^t_{2,k,1} \), \( f^t_{2,k,2} \), ... , \( f^t_{2,k,d} \)
- ... for each railway section
- Railway section rs - Period t: \( f^t_{rs,k,1} \), \( f^t_{rs,k,2} \), ... , \( f^t_{rs,k,d} \)

![Fig. 1 Structure of training datasets for period t](image-url)

**V. CASE STUDY AND NUMERICAL RESULTS**

The proposed approach has been tested on Line 1 of the Naples metro. This line (see Fig. 2) is long 18 km and has 18 stations; it connects high-density districts of Naples and is a crucial infrastructure for urban mobility.

Considering these characteristics, we have 18 or 34 turnstiles, if we do not divide the passengers in function of the direction or vice versa, and 34 mono-directional railway sections. The main features of the line are summarised in Table I.

The training datasets were obtained by simulating 2,500 OD matrices generated randomly; eliminating some of them because their results were not feasible (too many passengers compared to the actual capacity), we generated 2,279 training datasets. Among them, 2,229 were used for training the ANNs...
with the software MatLab, and 50 were used for verifying the results. We have tested six structures of ANNs for both cases: a) turnstiles at the entrance of the station (18 turnstiles), and b) turnstiles at the accesses to the platforms (34 turnstiles). Therefore, we have trained and tested 12 ANNs, as reported in Table II.

Table I Features of Line 1

| Feature                          | Value         |
|---------------------------------|---------------|
| Stations                        | 18            |
| Working day runs                | 241           |
| Convoy capacity (pax/convoy)    | 864           |
| Line length (km) (outward/return direction) | 18.8/18.6 |
| Headway (min)                   | 8-20          |

![Fig. 2 Line 1 route](image)

Table II Trained and tested ANNs

| Case | ANN     | Input nodes | Output nodes | Hidden layers | Neurons |
|------|---------|-------------|--------------|---------------|---------|
| a    | a_1_6   | 72          | 34           | 1             | 6       |
| a    | a_1_10  | 72          | 34           | 1             | 10      |
| a    | a_1_20  | 72          | 34           | 1             | 20      |
| a    | a_2_6   | 72          | 34           | 2             | 6/6     |
| a    | a_2_10  | 72          | 34           | 2             | 10/10   |
| a    | a_2_20  | 72          | 34           | 2             | 20/20   |
| b    | b_1_6   | 136         | 34           | 1             | 6       |
| b    | b_1_10  | 136         | 34           | 1             | 10      |
| b    | b_1_20  | 136         | 34           | 1             | 20      |
| b    | b_2_6   | 136         | 34           | 2             | 6/6     |
| b    | b_2_10  | 136         | 34           | 2             | 10/10   |
| b    | b_2_20  | 136         | 34           | 2             | 20/20   |

The training phase has required computing times from 30 seconds (case a_1_6) to 8 minutes (case b_2_20), with a PC i7-7700HQ, 2.80 GHz, RAM 16 GB. In Table III we report the best and worst coefficients of determination (R²) for each case, referring to the 50 datasets not used in the training phase, and the corresponding values of average and variance. The datasets for which the R² value is lower than 0.9, 0.8, 0.7 and 0.6 are reported in Table IV. In these tables, the best values for each ANN are underlined.

Examining the results reported in Tables III and IV, we may identify as best ANN structures the one with 1 hidden layer and 20 neurons for the case a), and 2 hidden layers and 10 neurons for the case b). The corresponding dispersion diagrams in the cases of best and worst R² are reported in Figs. 3 and 4.
Table III Coefficients of determination ($R^2$)

| ANN | Best     | Worst    | Average  | Variance |
|-----|----------|----------|----------|----------|
| a_1_6  | 0.9946   | 0.5487   | 0.7984   | 0.0124   |
| a_1_10 | 0.9941   | 0.4990   | 0.8108   | 0.0160   |
| a_1_20 | 0.9931   | 0.5613   | 0.8332   | 0.0157   |
| a_2_6  | 0.9916   | 0.5353   | 0.7949   | 0.0121   |
| a_2_10 | 0.9930   | 0.4638   | 0.8136   | 0.0175   |
| a_2_20 | 0.9933   | 0.5342   | 0.8244   | 0.0162   |
| b_1_6  | 0.9875   | 0.5460   | 0.8016   | 0.0129   |
| b_1_10 | 0.9905   | 0.3505   | 0.8115   | 0.0202   |
| b_1_20 | 0.9882   | 0.4226   | 0.8291   | 0.0160   |
| b_2_6  | 0.9785   | 0.5119   | 0.7775   | 0.0132   |
| b_2_10 | 0.9889   | 0.4935   | 0.8221   | 0.0132   |
| b_2_20 | 0.9852   | 0.4402   | 0.8075   | 0.0245   |

Table IV Analysis of $R^2$ values

| ANN | $R^2 < 0.9$ | $R^2 < 0.8$ | $R^2 < 0.7$ | $R^2 < 0.6$ |
|-----|-------------|-------------|-------------|-------------|
| a_1_6 | 41         | 25          | 9           | 4           |
| a_1_10 | 36         | 22          | 7           | 4           |
| a_1_20 | 33         | 18          | 7           | 3           |
| a_2_6  | 43         | 25          | 10          | 3           |
| a_2_10 | 35         | 21          | 8           | 4           |
| a_2_20 | 34         | 19          | 9           | 5           |
| b_1_6  | 37         | 24          | 10          | 2           |
| b_1_10 | 34         | 23          | 8           | 5           |
| b_1_20 | 32         | 19          | 8           | 3           |
| b_2_6  | 43         | 25          | 10          | 6           |
| b_2_10 | 35         | 20          | 7           | 2           |
| b_2_20 | 30         | 21          | 9           | 7           |

Percentage of datasets

| ANN | $82\%$      | $50\%$      | $18\%$      | $8\%$       |
|-----|-------------|-------------|-------------|-------------|
| a_1_6 | 72%         | 44%         | 14%         | 8%          |
| a_1_10 | 66%         | 36%         | 14%         | 6%          |
| a_1_20 | 66%         | 36%         | 14%         | 6%          |
| a_2_6  | 86%         | 50%         | 20%         | 6%          |
| a_2_10 | 70%         | 42%         | 16%         | 8%          |
| a_2_20 | 68%         | 38%         | 18%         | 10%         |

| ANN | $74\%$      | $48\%$      | $20\%$      | $4\%$       |
|-----|-------------|-------------|-------------|-------------|
| b_1_6 | 68%         | 46%         | 16%         | 10%         |
| b_1_10 | 64%         | 38%         | 16%         | 6%          |
| b_1_20 | 64%         | 38%         | 16%         | 6%          |
| b_2_6  | 86%         | 50%         | 20%         | 12%         |
| b_2_10 | 70%         | 40%         | 14%         | 4%          |
| b_2_20 | 60%         | 42%         | 18%         | 14%         |

VI. CONCLUSIONS AND RESEARCH PROSPECTS

In this paper, we proposed to use ANNs for forecasting passenger flows on railway sections of a metro line starting from counts at turnstiles. We considered two cases: turnstiles at station entrances and turnstiles at platform accesses. For both, we designed and trained some ANNs.
The results showed a good capacity of ANNs to forecast the loads on railway sections. Our analysis allowed us to identify the best ANN structure for each case. However, these results have to be considered only preliminary. Indeed, in future works we propose to: 1) test other ANN structures; 2) train a different ANN for each railway section, considering only some turnstile counts (the ones that forego the section under analysis); 3) propose deep learning approaches.

REFERENCES

[1] W. McCulloch, and W. Pitts, “A logical calculus of the ideas immanent in nervous activity,” Bulletin of Mathematical Biophysics, vol. 5, pp. 115-133, 1943.

[2] D.O. Hebb, The organization of behaviour. A neuropsychological theory. Wiley, New York, 1949.

[3] F. Rosenblatt, Principles of neurodynamics: perceptrons and the theory of brain mechanisms. Spartan Books, 1962.

[4] M. Minsky, and S. Papert, An Introduction to Computational Geometry. M.I.T. Press, Cambridge, Mass., 1969.

[5] T. Kohonen, “Self-organized formation of topologically correct feature maps,” Biological Cybernetics, vol. 43, pp.59-69, 1982.

[6] S. Grossberg, Neural Networks and Natural Intelligence. The Mit Press, Cambridge, 1988.

[7] M.L. Minsky, Theory of neural – analog reinforcement system and its application to the brain – model problem. Ph.D. Thesis, Princeton University, Princeton, NJ, 1954.

[8] J.J. Hopfield, “Neural networks and physical systems with emergent collective computational abilities,” in Proc. of the National Academy of Sciences of the U.S.A., vol. 79, 1982, pp. 2554-2558.

[9] J.S. Judd, Neural network modeling and connectionism. Neural network design and the complexity of learning. MIT Press, 1990.

[10] S. Haykin, Neural Networks A Comprehensive Foundation. McMaster University, Hamilton, Ontario, Canada, 1994.

[11] W.T. Miller, P.J. Werbos, and R.S. Sutton, Neural networks for control. MIT Press, 1995.

[12] R. Rojas, Neural Networks. A systematic introduction. Springer, 1996.
on sparse probe vehicle data, "traffic prediction using a binary neural network," UTSG, 2011.

Transportation Engineering prediction: bayesian combined neural network approach," W. Zheng, D.-H. Lee, and Q. Shi, "Short-term freeway traffic Transportation Research Record, vol. 1651, pp. 39-46, 1998.

volume forecasting using radial basis function neural network." B. Park, J.M. Carroll, and T.I. Urbank, "Short-term freeway traffic pp. 311-314, 1998.

A. Csikos, Z. J. Viharos, B. Kisk, T. Tettamanti, and I. Varga, "Traffic speed prediction method for urban networks an ANN approach," in Proc. of Models and Technologies for Intelligent Transportation Systems, Budapest, Hungary, 2015.

D. Park, and L.R. Rilett, “Forecasting Freeway Link Travel Times with a Multilayer Feedforward Neural Network," Computer-Aided Civil and Infrastructure Engineering, vol. 14, pp. 357–367, 1999.

R. Yasdi, "Prediction of road traffic using a neural network approach," Neural Computing & Applications, vol. 8, pp. 135-142, 1999.

R. Li, and H. Lu, “Combined neural network approach for short-term urban freeway traffic flow prediction," Advances in Neural Networks, Lecture Notes in Computer Science, vol. 5553, pp. 1017-1025, 2009.

Y. Gao, and S. Sun, “Multi-link traffic flow forecasting using neural networks," in Proc. of 2010 Sixth International Conference on Natural Computation, pp. 398-401, 2010.

J. Gao, Z. Leng, Y. Qin, Z. Ma, and X. Liu, “Short-term traffic flow forecasting model based on wavelet neural network," in Proc. of 2013 25th Chinese Control and Decision Conference, pp. 5081-5084, 2013.

C. Goves, “Short term traffic prediction on the UK motorway network using neural networks", in Proc. of European Transport Conference, 2015.

C. Ledoux, “An urban traffic flow model integrating neural networks," Transportation Research Part C, vol. 5, pp. 287-300, 1997.

L. Florio, and L. Mussone, “Neural network models for classification and forecasting of freeway traffic flow stability," in Transportation Systems: Theory and Application of advanced technology, Liu B., Blosseville J.M. (eds). pp. 773-784, 1995.

S. Lin, Y. Xi, and Y. Yang, "Short-Term Traffic Flow Forecasting Using Macroscopic Urban Traffic Network Model," in Proc. of Conference: 11th International IEEE Conference on Intelligent Transportation Systems, ITSC, 2008.

J. Zheng Zhu, J. Xin Cao, and Y. Zhu, “Traffic volume forecasting based on radial basis function neural network with the consideration of traffic flows at the adjacent intersections," Transportation Research Part C, vol. 47, pp. 139-154, 2014.

M. Gallo, and G. De Luca, “Spatial extension of road traffic sensor data with Artificial Neural Networks," Sensors, vol. 18, article id. 2640, 14 pages, 2018.

L. D’Acierno, M. Botte, and B. Montella, “Assumptions and simulation of passenger behaviour on rail platforms," Int. J. Transp. Dev. Integr., vol. 2, pp. 123-135, 2018.

[13] S.S. Haykin, Kalman filtering and neural networks. Wiley Online Library, 2001.
[14] C.C. Aggarwal, Neural Networks and Deep Learning. Springer, 2018.
[15] D. Baptista, and F. Morgado-Dias, “A survey of artificial neural network training tools," Neural Computing and Applications, vol. 23, pp. 609-615, 2013.
[16] F. Scarselli, and A.C. Tsoi, “Universal approximation using feedforward neural networks: a survey of some existing methods, and some new results," Neural Networks, vol. 11, pp. 15-37, 1998.
[17] S. Timotheou, “The random neural network: a survey," The Computer Library, vol. 53, pp. 251-267, 2010.
[18] G. Huang, G.-B. Huang, S. Song, and K. You, “Trends in extreme learning machines: A review," Neural Networks, vol. 61, pp. 32-48, 2015.
[19] J. Schmidhuber, “Deep learning in neural networks: An overview," Neural Networks, 2015, pp. 85-117, vol. 61.
[20] W. Liu, Z. Wang, X. Liu, N. Zeng, Y. Liu, and F.E. Ahsaadi, “A survey of deep neural network architectures and their applications," Neurocomputing, vol. 234, pp. 11-26, 2017.
[21] X. Yao, “A review of evolutionary artificial neural networks," International Journal of Intelligent Systems, vol. 8, pp. 539-567, 1993.
[22] G. De Luca, M. Gallo, “Artificial Neural Networks for forecasting user flows in transportation networks: literature review, limits, potentialities and open challenges," in Proc. of 2017 5th IEEE International Conference on Models and Technologies for Intelligent Transportation Systems, Naples, pp. 919-923, 2017.
[23] E.I. Vlahogianni, M.G. Karlaftis, and J.C. Golias, “Short-term traffic forecasting: where we are and where we’re going," Transportation Research Part C, vol. 43, pp. 3-19, 2014.
[24] S. Oh, Y.J. Byon, K. Jang, and H. Yeo, “Short-term travel-time prediction on highway: a review of the data-driven approach," Transport Reviews, vol. 35 , pp. 4-32, 2015.
[25] H.R. Kirby, S.M. Watson, and S. Dougherty, "Should we use neural networks or statistical models for short-term motorway traffic forecasting?", International Journal of Forecasting, vol. 13, pp. 43-50, 1997.
[26] B.L. Smith, and M.J. Demetsky, “Short-term traffic flow prediction: Neural network approach," Transportation Research Record, vol. 1453, pp. 98-104, 1994.
[27] H.J. Zhang, S.G. Ritchie, and Z.P. Lo, “Macroscopic modeling of freeway traffic using an artificial neural network," Transportation Research Record, vol. 1588, pp. 110–119, 1997.
[28] M.S. Dougherty, and H.C. Kirby, “The use of neural networks to recognize and predict traffic congestion," Traffic Eng. Control, vol. 346, pp. 311-314, 1998.
[29] B. Park, J.M. Carroll, and T.I. Urbank, “Short-term freeway traffic volume forecasting using radial basis function neural network." Transportation Research Record, vol. 1651, pp. 39-46, 1998.
[30] W. Zheng, D.-H. Lee, and Q. Shi, “Short-term freeway traffic prediction: bayesian combined neural network approach," Journal of Transportation Engineering, vol. 132, pp. 114-121, 2006.
[31] V. Hodge, J. Austin, R. Krishnan, J. Polak, and T. Jackson, “Short-term traffic prediction using a binary neural network," UTSG, 2011.
[32] K. Kumar, M. Parida, and V.K. Katiyar, “Short term traffic flow prediction for a non urban highway using Artificial Neural Network,” Procedia - Social and Behavioral Sciences, vol. 104, pp. 755-764, 2013.
[33] F. Zheng, and H. Van Zuylen, “Urban link travel time estimation based on sparse probe vehicle data," Transportation Research Part C, vol. 31, pp. 145-157, 2013.