The Current Scientific Stage of The Instruments and Methods Needed for an Efficient Traffic Management System Based on AI

Florin ANDREESCU
Bucharest University of Economic Studies, Romania
andreescuflorin19@stud.ase.ro

An efficient traffic management system based on AI has to address several main objectives like significant reduction of overall congestion levels in the city and dynamic pathways generation for emergencies and government using intelligent informed and self-unlocking intersections. Such a system must consider also future trends and challenges like autonomous vehicles, automated courier services, and flying vehicles in the city. This type of management is a complex objective that aims at both the optimization of flows and the optimal response to disturbances and crisis situations. This study investigates the current scientific stage of the instruments and methods needed for architecture, design, construction, and implementation of an informatics management system based on AI able to solve such a complex goal.
Keywords: *Informatics Management System, AI, Traffic, Congestion, Emergencies, Machine Learning*
DOI: 10.24818/issn14531305/26.1.2022.05

1 Introduction

The long-term objective of the research is to define an Information Management System for Urban Traffic Management (IMSUTM).

The main goals are the significant reduction of general congestion levels in the city using intelligent informed and self-unlocking intersections and the dynamic pathways generation for emergencies and government with minimal human agents' intervention.

There is also a list of perspective goals that will evolve in time, which includes integrations with autonomous vehicles, automated courier services, and openness to 3D to manage city aerial pathways for flying vehicles in the city.

This is a complex objective that aims at both the optimization of flows and the optimal response to disturbances and crisis situations. It involves the cooperation of several intelligent systems with roles like Real-time data collection, IoT, Prediction, Decision/Execution, and Interaction with Actuators in intersections.

An important component of this ensemble is the Predictive Core of traffic levels which has to offer valuable information for the decision levels and at the city intersections level.

Complex historical data over several years are needed to train predictive machine learning (ML) models.

Because in most cases historical information is not fully available, is necessary to create a Historical Data Generator that combines simulated data with real data to generate data that reproduces the behavior of the city as much as possible.

The paper is structured as follows. Section 2 presents the main ideas related to the Historical Data Generator and the last section will present the approaches related to the key components of the IMSUTM Predictive Core of traffic levels. Section 3 presents the approaches found related to similar systems currently in existence in the world and related to the key components of the Historical Data Generator. The paper ends with conclusions and future work.

2. Historical Data Generator

A key component of the Historical Data Generator is the City Traffic Simulator. Starting from a set of socio-behavioral matrices which describes the city's "habits", it must generate for each day all the routes made by each person within the stable population of the city completed with the set of routes made in transit by people outside the city (entrances-exits from the city).

Each route generated is composed of a chain

of roads and breaks and for most of the population, this chain ends at the starting point (home).
Each generated road is accompanied by attributes such as type of transport (pedestrian, car, public transport, bicycle, or scooter). Depending on the start time and the weather, it will pass through regions of the city that will slow it down variably depending on the congestion at that time or the weather conditions.
A data set used for training the predictor models must be the result of a simulation for at least 5 years. This dataset is mixed gradually with real data provided by trusted sources (authorities, research institutes, and any other informatic systems capable to provide relevant information).
At the start, the main real data source is the historical dataset about the weather in the last 5 years.
Finally, the result will be a labeled synthetic data set ready to be used to train the prediction ML models.
The key advantage of this component is the possibility to “inject” patterns and specific correlations during the generation process. For example, correlated with the time are the student vacations and the holiday seasons. Unexpected patterns are not necessarily correlated with the time factor. They could be contests, championships, festivals, parades, or even pandemic and disaster patterns.

3 Approaches related to the key components

A real IMSUTM will contain a chain of components that will cover activities like Real-time data collection, IoT, Prediction, Decision/Execution, and Interaction with Actuators in intersections.
The next section will present an overview of the similar existing systems in the world. Further sections will show the approaches related to the main components involved in the Historical Data Generator.

3.1 Approaches to advanced traffic management

In cities where the number of vehicles is constantly growing faster than the traffic infrastructure available to support them, congestion is a difficult problem to solve and becomes even more serious in the event of car accidents.
This problem affects many aspects of modern society, including economic development, road accidents, increasing greenhouse gas emissions, time spent, and damage to health. In this context, modern societies can rely on the traffic management system to minimize traffic congestion and its negative effects.
According to [1], traffic management systems consist of a set of application and management tools to improve the efficiency of general traffic and the safety of transport systems.
In addition, to overcome such a problem, the traffic management system collects information from heterogeneous sources, exploits such information to identify hazards that could degrade traffic efficiency, and then provide services to control it.
In [1] are presented a classification, review, challenges, and perspectives for implementing a traffic management system
According to Quartz India, India's largest cities lose about $ 22 billion annually due to fuel consumption and lost productivity. The Delhi Police Department is in the process of implementing an intelligent traffic system, with automated traffic signaling as a key component. 7,500 cameras with multidirectional sensors were used to count the volume of traffic on all Delphi arteries, based on the patterns detected in the image. "The goal is to keep manual intervention to a minimum. Intelligent traffic lights, artificial intelligence to determine the optimal course of traffic"- Dependra Pathak, Commissioner of the Traffic Department of the Delphi Police [2].
The areas of application of AI for solving traffic and transport problems are summarized in [3].
Book [4] describes a concept similar to the General Objective of this project and details the role of the components involved, especially the prediction part based on artificial intelligence and the relationship with

traffic actuators. The spectacular idea highlighted is that smart systems need to solve traffic problems by simulating the current behavior of the Traffic Police.
Among the existing products on the market that aim to address the problem of traffic, it stands out:
Complete AI-Based Advanced Traffic Management System, Telegra. topXview ™ is used in over 100 traffic monitoring agencies worldwide. topXviewTM uses advanced artificial intelligence techniques to analyze huge amounts of data that are collected from various sources, video analysis, and offers advanced decision systems that use AI to predict and manage traffic and road accidents [5].
Ericsson promotes traffic management using a dedicated architecture - MECA - Multipurpose Enhanced Cognitive Architecture. This is represented by a network of cognitive managers, each responsible for a dedicated area of the city, but who can communicate with each other to achieve global traffic optimization [6].
The idea of automatic traffic management for crowded cities has preoccupied both the authorities and the private segment for many years.
There have been various approaches in the world based on the top technologies of each period.
An overview of advanced traffic management can be found in an article on the Cornell University website [7]. There is also a classification of data sources for such systems. It is emphasized that automated methods can provide more information by supplementing it with data sources provided by GPS and GSM systems.
In our country, the UTI company has been using the Urban Traffic Management (Utopia) product for several years in a continuous effort to improve the traffic conditions in Bucharest and in other cities in the country.
UTI Group is a Romanian company that offers integrated traffic management solutions. The product offers real-time traffic monitoring, intelligent traffic lights, and transportation time measurement. However, the information provided by the company does not indicate the use of artificial intelligence systems to provide these services. The product is implemented in SZCZECIN (Poland), Galati, Iasi, Deva, Arad, Bucharest. UTI operates as a systems integrator, and the traffic management component is called Utopia and was provided by MIZAR. The estimated improvement in traffic conditions is 15-18% [8]
Statistics on congestion levels for Bucharest for each year can be tracked in TomTom reports [9].

3.2 Approaches to data collection

An interesting trend is the use of IoT to measure traffic in a certain area. IoT refers to sensors that can automatically transmit traffic information. They can be specialized or use GPS and GSM systems (GPS position of the car or phone). Dedicated sensors usually transmit information with a certain frequency to a command-and-control center - usually 60 seconds. This time may be too long for crowded areas, but it can be optimized by using a predictor [10].
One way to collect traffic data is presented in Chapter 3 from [11]. Cross-sectional data is collected by cross-road sensors and is considered a very reliable way of collecting data.
In [12] is presented the importance of License Plate Recognition systems. They are important and very reliable data sources for Systems that aim to analyze and manage traffic.
Another important source of data relates to weather data. There are official sources such as the INMH Archive or Meteoblue [13] for data from recent years. INMH has also to provide the present and estimated data for the next hours/days.

3.3 Simulation approaches

An interesting classification is described in [14]. The simulators are grouped into classes considering the level of resolution they refer to (Macroscopic, Mesoscopic, and Microscopic).
The simulators that cover the whole spectrum

are called Multimodal. The current Traffic Generator aims at the Macroscopic level and in the future, the other levels will be approached according to needs.
A general classification of various similar traffic simulation products in the world can be found in [15].
In the area of multimodal simulators, we find products that go to advanced degrees of detail such as PTV VISSIM [16] and SimWalk [17].
In the area of development frameworks for simulators, MATSim [18] as well as SUMO stand out.
MATSim is a Java Open Source framework to implement large-scale agent-based transport simulations. It was successfully applied to various domains like road transport public transport, freight transport, regional evacuation. The project was developed in similarity with the computational physics simulators, considering that behavior-oriented simulations with 10 million travelers should be possible in travel behavior research.
The Historical Data Generator in this research will be defined starting from the same principle with the augmented functionality regarding the requirement to have the possibility to inject custom patterns and correlations.
The SUMO simulator is a software tool developed by the Institute of Transport Systems of the German Aerospace Center (DLR). It is a microscopic and continuous platform for simulating road traffic and is widely used in urban traffic research.
SUMO is a free and open traffic simulation suite, available since 2001. SUMO allows the modeling of intermodal traffic systems, including road vehicles, public transport and pedestrians. Includes SUMO is a variety of support tools that handle tasks such as finding routes, viewing, importing networks, and calculating emissions. SUMO can be enhanced with custom models and offers various APIs to remotely control the simulation [19].
The book [20] describes the use of SUMO in conjunction with a relational learning process based on the phased formulation of goals and metrics as well as through continuous monitoring of results.

3.4 Prediction and classification approaches

IMSUTM will include a Prediction component named The Predictive Core of traffic levels. This component has to predict future traffic levels in each intersection of the city in order to inform the decision levels and the Intersection Component. Deep Learning Models training is required to achieve the Prediction goal.
In the same way, the Intersection Component will consist of a set of Deep Learning Models trained specifically for each intersection of the city. Each one of these models will be able to take local decisions in real-time to adequately command the actuators in terrain.
In [21] is presented the relationships between AI, ML, deep learning, and data science, the description of ML types and the structure of ANN (Artificial Neural Networks) (Figure 1).

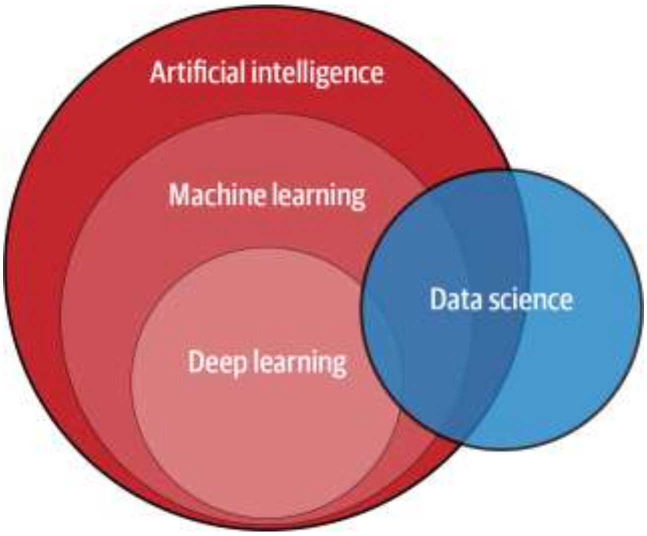


Fig. 1. The relationships between AI, ML, deep learning and data science [21]

The large domain of **Artificial intelligence** is the study of the development of machines' abilities to perform complex operations that usually require human intelligence which include **Machine learning** (ML) refers to machine learning processes to make better and better decisions, **Deep learning** is a sub-branch of ML that involves the study of algorithms based on neural networks and a part of **Data Science** which is an interdisciplinary science that uses scientific methods, processes and, systems to extract knowledge from data in whatever form they are accessible.

The main classification of **Machine Learning types** is shown in Figure 2.

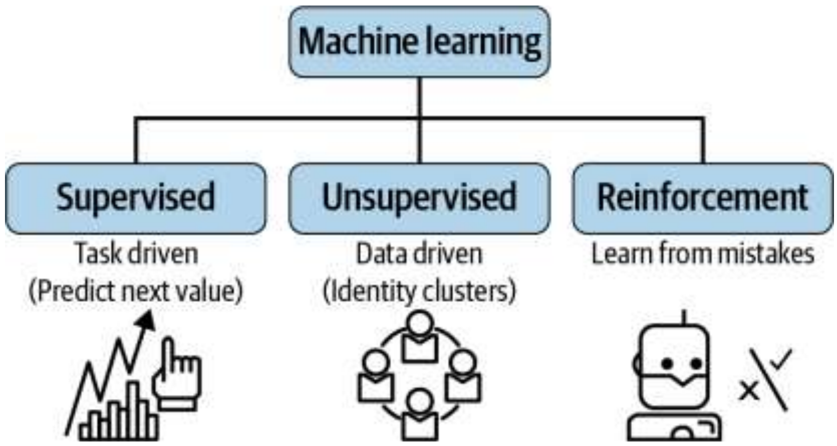


Fig. 2. ML types [21]

Supervised learning - training models based on data sets already labeled. It includes two basic types: **classification** whose purpose of the classification is to predict the classes of objects based on previous observations and **regression** which is able to predict the output values in a continuous mode depending on the set of input values.

Unsupervised learning represents the analysis of previously untagged data. It includes two basic types: **Dimensionality reduction** which studies the possibility of reducing the number of input variables to maintain the level of performance and **Clustering** which studies the discovery of hidden structures in data (natural groups of data).

Reinforcement learning (RL) studies algorithms for learning from experience through reward-punishment mechanisms.

Another interesting domain is **Natural**

language processing (NLP) is a branch of AI that refers to the ability of machines to understand people's natural language.
Another important concept is represented by the Artificial Neural Networks. The next paragraph presents the architecture of neural networks as well as the mechanisms related to the learning process. Thus, the ANN architecture is composed of neurons, levels, and weights.
A neuron receives a set of input data and applies a weighted amount based on weights to each input (Figure 3).

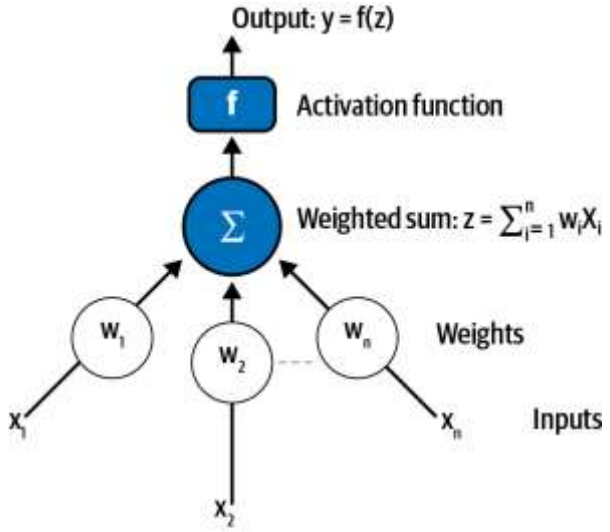


Fig. 3. Neuron structure [21]

The resulting weighted sum is passed to an activation function (f).
The Activation functions (Figure 4) could be: Linear, Sigmoid, Tanh or ReLU.
Linear (identity) function is represented by the equation of a straight line like

$$f(x) = mx + c.$$

Here activation is proportional to the input. The output can have values between –infinite to +infinite.
Sigmoid function (logistic activation function) is represented by the mathematical equation:

$$f(x) = 1/(1 + e^{x}).$$

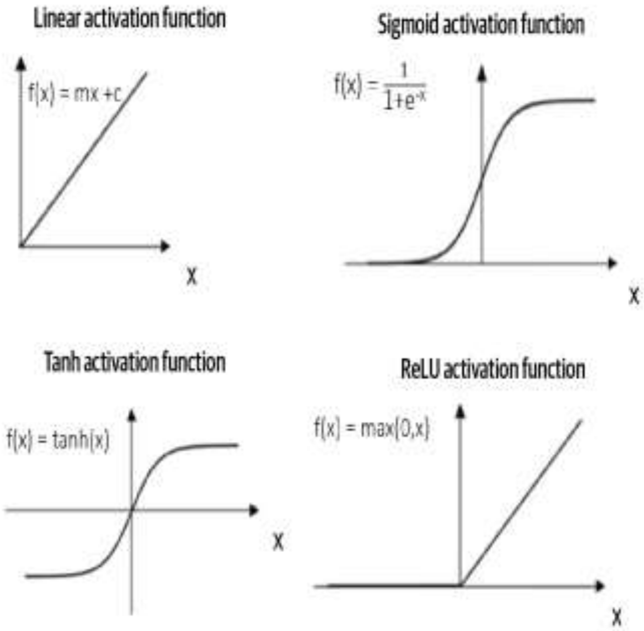
The output can have values between 0 and 1.
Tanh function is represented by the mathematical equation:

$$Tanh(x) = 2Sigmoid(2x) – 1,$$

where Sigmoid represents the sigmoid function. The output of this function can have values between –1 to 1.
ReLU (Rectified Linear Unit) function is represented as

$$f(x) = max(x, 0).$$

It is the most used function because of its simplicity.



. **Fig. 4.** Graphs of Activation functions [21]

In the ANN, several levels are defined that contain neurons. These can be input, output and hidden, as presented in Figure 5.

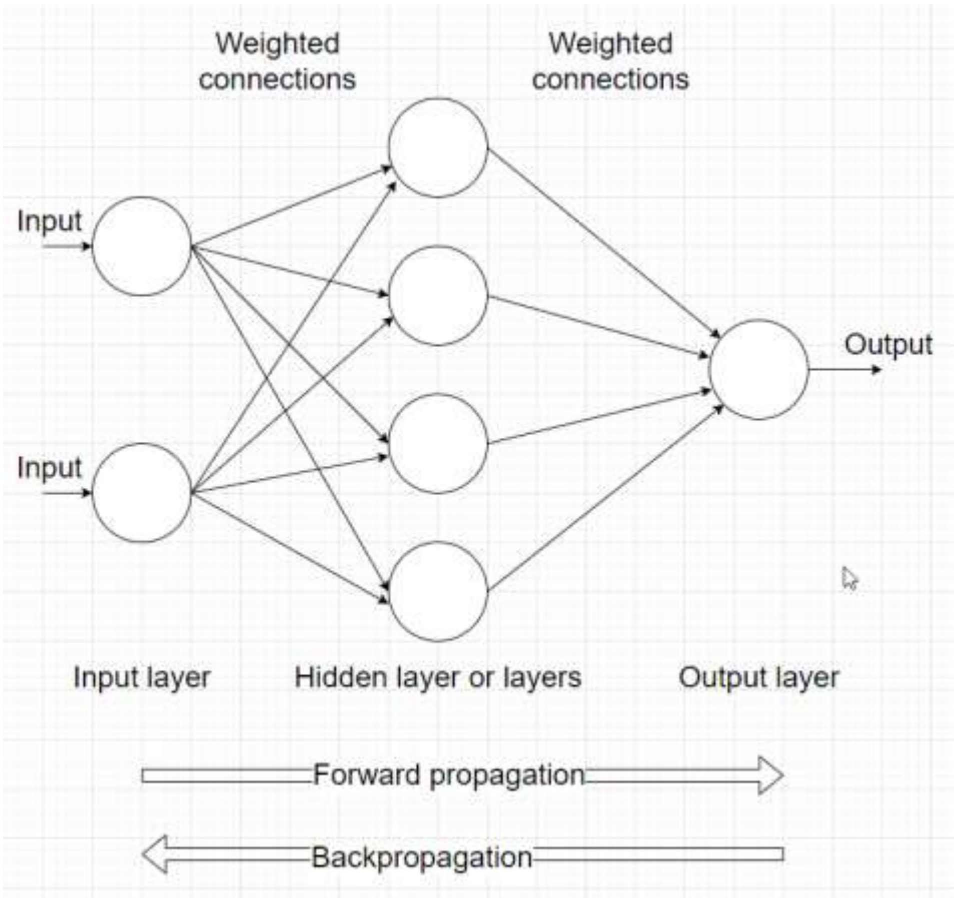


Fig. 5. Layers

Learning involves the iterative process of recalibrating all weights in the network.
Forward propagation - apply a set of input data and calculate the output (predicted value)
Backpropagation - calculates the loss or cost (the difference between the output obtained and what it should have been).
If the real (labeled) output is Y and the obtained value is Y' then (Y-Y') is converted into the loss function J(w), where **w** are the weights in the entire neural network. The goal is to optimize the loss function to minimize losses.
The optimization method used is called gradient descent (Figure 6). This is based on applying small steps in the direction of the negative gradient until the minimum value is reached.

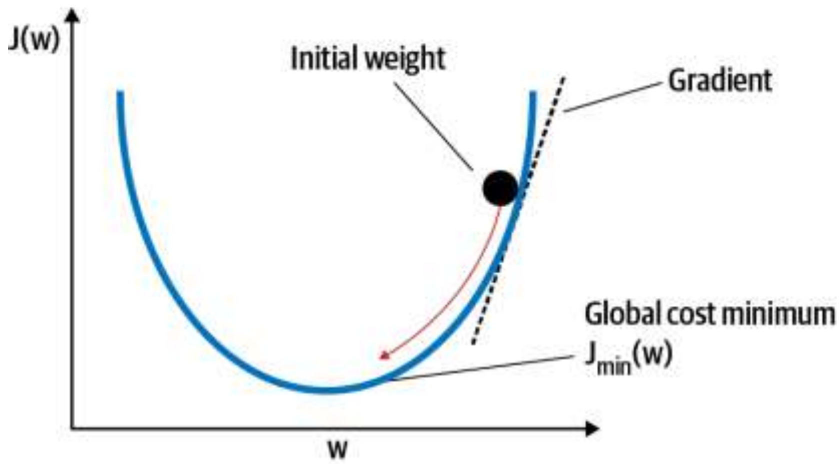


Fig. 6. Gradient descent [21]

The process of backpropagation found the loss (error) in reverse on the layers (from output to reinforcement) to recalculate the weights on each layer.
Hyperparameters are fixed variables during the learning process. They have an important influence on the learning process and therefore, its regulation is done by iteratively modifying the Hyperparameters and resuming the learning process. Among them the most frequently used are:

a. The number of levels and the number of neurons on them
b. Learning rate
c. Activation functions (Linear, Sigmoid, Tanh, ReLU), Figure 4**.**
d. Cost functions:
   - Mean squared error (MSE) – quadratic mean of error - used for regressions
   - Cross-entropy (or log loss) – used for classification given that the output is a probability with values between 0 and 1.
e. Epochs - the number of complete passes through the learning dataset
f. Batch size – the number of examples in the training set used in a complete forward / backward cycle.

Optimizers - these update the values of the weights in the network in order to minimize the loss function. The most used optimizers are: Momentum, AdaGrad (Adaptive Gradient Algorithm), RMSProp (Root Mean Square Propagation), Adam (Adaptive Moment Estimation).
An in-depth presentation of AI algorithms and their applications in the real live could be found in [22]. In addition to the domain of Supervised Learning, the book cover also the domains of Unsupervised Learning and Reinforcement Learning and explores how to make rational decisions regarding actions to be taken in a context where the state variables on which a decision is based are themselves uncertain.
The study [23] analyzes human position games and artificial intelligence algorithms in

decision making.
Because both humans and intelligent algorithms suffer from biases, the collaboration between them must be carefully studied and regulated. Human operators integrated into the same decision loop have a very limited ability to correct any "biases" or errors already existing in the trained models.
The book [24] offers comprehensive support to use of TensorFlow for creating, training, and using machine learning models.
For to cover all the goals required in the traffic management system is important to study also emerging domains presented in [25] like Neural Network Toolkits, Autoencoders, and Recommender Systems.

4. Conclusions

The research proposes a way to solve the problem of the urban traffic estimator using technologies for data ingestion, processing, and AI algorithms.
The real-time urban traffic estimator is a very important component in the process of obtaining a Universal Urban Traffic Regulator necessary in the conditions of accelerated agglomeration of cities.
The promising results obtained from the studies both on the generation and simulation side and on the training side of ML models for prediction, decision, and control allow the transition to the experimental implementation phase and evaluation of results for the components necessary to achieve the targeted Information Management System for Urban Traffic Management.

References

[1] C. A. B. R. S. Y. E. A. D. E. R. M. L. A. V. Allan M de Souza, „Traffic management systems: A classification, review, challenges, and future perspectives,” *International Journal of Distributed Sensor Networks,* vol. 13, nr. 10.1177/1550147716683612, p. 04, 2017.
[2] V. Kumar, „Artificial Intelligence: Application in traffic management system,” 11 Oct 2018. [Interactive]. Available: https://medium.com/@wrandomwriter/artificial-intelligence-application-in-traffic-management-system-9b5e3d7e620c. [Accessed 20 06 2019].
[3] R. K. M. P. Rndr. Šusteková D. PhD., „How Is the Artificial Intelligence Used In Applications For Traffic Management,” in *Scientific Proceedings Xxiii International Scientific-Technical Conference "Trans & Motauto ’15" Year Xxiii, Volume 3*, 2015.
[4] B. L. Sun, Application and Study on Artificial Intelligence Technology in Traffic Signal Control System, Springer, Berlin, Heidelberg ISBN: 978-3-642-23222-0, 2011.
[5] Telegra, „Complete AI Based Advanced Traffic Management System,” 2019. [Interactive]. Available: https://www.telegra-europe.com/solutions/solution-664-complete-ai-based-advanced-traffic-management-system. [Accessed 10 06 2019].
[6] Ericsson, „Avoiding gridlock using cognitive architecture and smart traffic management systems,” Ericsson, 2019. [Interactive]. Available: https://www.ericsson.com/en/blog/2019/1/avoiding-gridlock-using-cognitive-architecture-and-smart-traffic-management-systems. [Accessed 15 02 2019].
[7] D. G. M. Shahgholian, „Advanced Traffic Management Systems: An Overview and A Development Strategy,” arXiv funded by Cornell University, 2018.
[8] Uti, „Bucharest, Romania – Adaptive traffic management system,” UTI, 2019. [Interactive]. Available: https://www.uti.eu.com/business-lines/intelligent-transportation-solutions/urban-traffic-management/portfolio/bucharest-romania-adaptive-traffic-management-system/. [Accessed 05 06 2019].
[9] TomTom, „Bucharest in the Traffic Index,” TomTom, 2018. [Interactive]. Available: https://www.tomtom.com/en_gb/traffic-index/bucharest-traffic#statistics. [Accessed 02 06 2019].

[10] M. V. S. A. K. N. R. A. S. J. P. Senthil Kumar Janahan, „IoT based smart traffic signal monitoring system," *International Journal of Engineering & Technology*, vol. 7, nr. (2.21) (2018) 309-312, 2018.
[11] A. K. M. Treiber, Traffic Flow Dynamics, Springer, 2013.
[12] L. J. X. Z. V.E. Balas, „Information Technology and Intelligent Transportation Systems," in *Proceedings of the 2nd International Conference on Information Technology and Intelligent Transportation Systems*, Xi'an, China, 2017.
[13] Meteoblue, „Descărcare istoric meteo București," Meteoblue, 2019. [Interactive]. Available: https://www.meteoblue.com/ro/vreme/archive/export/bucure%C8%99ti_rom%C3%A2nia_683506. [Accessed 11 06 2019].
[14] Z. Y. R. G. a. F. D. Christoph Sommer, „Simulating the Influence of IVC on Road Traffic," *IEEEE,* vol. 08, nr. 978-1-4244-2219-7, p. 1, 2008.
[15] A. M. O. S. Andreas Pell, „Trends in Real-time Traffic Simulation," in *Transportation Research Procedia Volume 25, 2017, Pages 1477-1484*, Shanghai, 2016.
[16] Ptv Vissim, „The world's leading multimodal simulation software," Ptv Vissim, 2019. [Interactive]. Available: https://www.ptvgroup.com/en/solutions/products/ptv-vissim/areas-of-application/traffic-flow-simulation/. [Accessed 01 06 2019].
[17] SimWalk, „Make the Best of Your Public Transport Infrastructure," SimWalk, 2019. [Interactive]. Available: https://www.simwalk.com/simwalk_transport/index.html. [Accessed 10 05 2019].
[18] MATSim, „MatSim is an open-source framework...," MATSim, 2019. [Interactive]. Available: https://www.matsim.org/. [Accessed 03 06 2019].
[19] Dlr - Institute of Transportation Systems, „Eclipse Sumo – Simulation of Urban Mobility," DLR - Institute of Transportation Systems, 2019. [Interactive]. Available: https://www.dlr.de/ts/en/desktopdefault.aspx/tabid-9883/16931_read-41000/. [Accessed 23 03 2019].
[20] S. F. D. B. Alberto Pozanco, „Urban Traffic Control Assisted by AI Planning and Relational Learning," în *ATT@IJCAI*, 2016.
[21] S. P. a. B. L. Hariom Tatsat, Machine Learning and Data Science Blueprints for Finance, Sebastopol, CA 95472: O'Reilly Media, Inc., 2021.
[22] J.-A. Goulet, Probabilistic machine learning for civil engineers, Cambridge, Massachusetts: The MIT Press, 2020.
[23] J. W. Michelle Vacarro, „The Effects of Mixing Machine Learning and Human Judgment," *acmqueue,* p. 22, July-August 2019.
[24] L. Moroney, AI and Machine Learning for Coders, Sebastopol, CA 95472: O'Reilly Media, Inc., 2021.
[25] A. F. Vermeulen, Industrial Machine Learning, West Kilbride, UK, 2020.

Florin ANDREESCU has graduated the Faculty of Computer Science and Automation in The Polytechnic Institute Bucharest (UPB), Domain: Computers/Software Engineering. Currently he is PhD student at Economic Informatics Doctoral School - Bucharest University of Economic Studies (ASE), Domains: Management Information Systems, Artificial Intelligence, and Simulators. He works at Totalsoft as a senior developer and is in charge of an important project at Orange Romania. He has an extensive experience in complex projects including projects for research, industry, governmental and private companies. He was trained in Google and Azure cloud platforms and in recent projects, he worked in domains like Scalable Machine Learning and Video Streaming in Google cloud, Pub-Sub, Data Flow, Big Query, Auto ML, Vision ML integration, ERP/ORM integration, Data Warehouse, ETL and ML pipelines.