PCA modeling of extraction and selection of variables influencing LTE network delay in urban mobility conditions

Mirko Stojčić[1], Milorad K. Banjanin[2], Milan Vasiljević[3], Aleksandar Stjepanović[1], Zoran Ćurguz[1]

[1] *Faculty of Transport and Traffic Engineering Doboj, University of East Sarajevo, Vojvode Mišića 52, 74000 Doboj, Bosnia and Herzegovina*

[2] *Faculty of Philosophy Pale, University of East Sarajevo, Alekse Šantića 1, 71420 Pale, Bosnia and Herzegovina; University of Novi Sad, Faculty of Technical Sciences, Trg Dositeja Obradovića 6, 21000 Novi Sad, Serbia*

[3] *Faculty of Philosophy Pale, University of East Sarajevo, Alekse Šantića 1, 71420 Pale, Bosnia and Herzegovina*

Abstract

In addition to network throughput, which defines the number of bits per second that can be transmitted, another primary measure of network performance is latency, or delay, which defines the time spent sending one bit from one computer to another in the network. In this paper, the main goal of the research is to model the efficient management of delay performance in a 4G Long Term Evolution (LTE) network on a defined section of the trunk road with the assumption that key performance indicators (KPI) can be effectively analyzed with the created predictive classification models. These models also enable the assessment and prediction of the fulfillment of sustainable urban mobility plans. Using Principal Component Analysis (PCA), the space of 17 input variables is reduced to four extracted components. Several models based on different machine learning techniques are created using the automatic modeling method, and the final solution is selected according to the criterion of maximum classification accuracy, interpretability, and complexity. A classification model based on Logistic Regression was chosen as the final solution.

***Keywords**: Predictive classification models, PCA, network delay, LTE network, KPI, urban mobility*

1 Introduction

The trend of constant growth in demand for various 4G, and today also 5G and 6G services, obliges telecommunication providers to carefully plan and predict the required level of Quality of Service (QoS) to meet user needs in the future [1]. This problem is largely solved by continuously finding ways to efficiently share network radio resources with scheduling algorithms to meet different QoS profiles. The two basic Key Performance Indicators (KPIs) for Real-Time (RT) services, which today make up an increasing part of the total traffic generated in the network, are throughput and delay [1]. End-to-end delay or, in general, delay, can be defined as the time required for a data packet to be transmitted through the network from the source node to the destination node. It can be represented mathematically by the sum of transmission delay, signal propagation delay, and packet queuing delay. Therefore, the delay depends on many influential factors such as the path length from the source to the destination, intermediate hops in the network, interference, protocols, signal strength, medium properties, medium access control, etc. [2]. The precise formulation of a mathematical model for packet delay prediction is therefore very difficult. In recent years, the solution has been found in the methods and techniques of machine learning, i.e., in Data-Driven models. Contemporary trends are directed towards the creation and prioritization of interpretable models, ahead of complex models that, as a rule, provide better performance. Excessive network delay can significantly negatively affect the

Corresponding author: *Mirko Stojčić (mirko.stojcic@sf.ues.rs.ba)*
© 2023 The Author(s). This work is licensed under a Creative Commons Attribution 4.0 International License.

Quality of Experience (QoE) of users of telecommunication services, especially in conditions of high mobility in cities [2, 3]. This influence is expressed by measures of rigidity that services cause in users: Interpersonal (R1), Behavioral (R2), Structural - longing for structure and coping with the lack of structure (R3), Prospective anxiety (R4), Inhibitory rigidity (R5) [2, 3].

Modern, and especially mobile networks of the next generation, are characterized by a very large amount and variety of network variables - KPI. They significantly complicate the process of their predictive modeling. In the current research, an efficient selection of a smaller set of KPIs, i.e. their influencing factors, using Dimensionality Reduction techniques is proposed as a solution, which includes two approaches: Feature Selection and Feature Extraction [4]. Compared to Feature Selection, Feature Extraction techniques can retain a larger amount of information from the initial, complete data set. In modeling network performance, two families of Feature Extraction techniques are of particular importance:

1) Component Analysis, the most popular of which are Principal Component Analysis (PCA), kernel PCA (kPCA), and Independent Component Analysis (ICA);
2) Manifold Learning, of which the most commonly used are Locally-linear embedding (LLE) and Spectral embedding (SE) [4].

Today, mobile users and providers are witnessing the growing use of multimedia real-time applications. This is of particular importance for researchers, especially in urban areas characterized by high density of users and high mobility of users. In such circumstances, network delay is a decisive factor for a satisfactory level of QoE and QoS. Therefore, the main purpose and reason of this research is to contribute to the predictive modeling of delay in urban mobility conditions. The research is aimed at end-to-end delay (as a dependent variable) in the 4G LTE network in the area of the section of the M17 trunk road in Republic of Srpska, Bosnia and Herzegovina. The task is to create a machine learning model that, based on given input variables, can classify the delay. In this paper, delay is represented by a discrete scale with three levels. The created models must have sufficiently high accuracy. To reduce the dimensionality of the space of input variables, PCA is applied over the original set of 17 research variables as one of the most common linear feature extraction techniques. In the final step, a comparison of the classification results is made for the complete set and the set of extracted input variables. Based on that, the final solution is chosen with the assumption that the created models can be universally used for all Real-Time services.

Predictive delay modeling is the subject of numerous published studies. Papers [2], and [5] – [10] explore models based on different machine learning techniques, in different network environments and in combination with other approaches. A more detailed analysis of these studies is given in Section 2.

The research on network delay in this paper is of great importance from the aspect of sustainable urban mobility. In addition to the observed section being part of one of the most frequent road routes in Republic of Srpska and Bosnia and Herzegovina, urban and suburban transport of a large number of passengers takes place along it. At the same time, it is considered that passengers are also users of various telecommunication services, most of which are very sensitive to delays. In such circumstances, predictive delay models can be used to predict and assess the degree of fulfillment of the Sustainable Urban Mobility Plans (SUMP) in the domain of quality of the urban environment and quality of life.

As the main contributions presented in this paper, the following can be singled out:

- To the best of our knowledge this is the first research related to the modeling of delays in the LTE network in the observed area of the road section in the context of urban mobility;
- Several machine learning models were trained and tested and the best one was selected from among them;
- Created classification models have an overall accuracy of over 90% with a PCA reduced set of input variables;
- The selection of the final model, apart from classification performance, was made according to the criteria of interpretability and complexity;
- The observed traffic section geographically extends approximately parallel to and near the section of the pan-European corridor Vc. It gives additional importance and weight to this research because the results are applicable in that area as well;
- The results presented in this research can be of practical importance when designing or

expanding the network in the observed geographic area.

The structure of the paper consists of five sections. After the introductory part, the second section provides an overview of relevant published research. The third section describes the steps of the methodological research procedure, as well as the materials used. The main focus is on the fourth section, where the results of the research and discussion are presented. Concluding considerations are given in the last, fifth section. At the end of the paper is a list of used references.

2 Review of Relevant Published Research

To date, more research has been published related to predictive modeling of delays in 4G LTE networks. In paper [2], the goal is to evaluate the applicability and ability of an artificial neural network for end-to-end packet delay prediction in a mobile ad-hoc network environment. As input variables, the authors use the length of the path and the average number of hops between the source and the destination. Round-trip time (RTT) delay is the subject of research in [5]. The problem being solved is the prediction of changes in the value of this variable in mobile broadband networks. Four classification models based on machine learning were created in the paper, and the binary ensemble model showed the best performance. In the paper [6], the authors present an approach based on machine learning, which, thanks to data from real mobile networks, performs delay prediction. A large data set with more than 238 million delay measurements for three different commercial mobile operators is considered. A delay prediction framework that integrates machine learning and statistical approaches is defined in [7]. The first created model is based on Long Short-Term Memory (LSTM), and the second is based on a combination of Epanechnikov Kernel and moving average function. In paper [8], a low-complexity algorithm for predictive resource allocation for delay-sensitive Machine to Machine (M2M) services in the LTE uplink direction is specified. One of the key issues in 4G networks is how to efficiently use radio resources. The proposal of the new LTE packet scheduler mechanism and Packet Prediction Mechanism (PPM) was given by the author's team in the research [9]. Paper [10] deals with the topic related to the Industrial Internet of Things (IIoT), in which predictive delay models are created to evaluate network performance.

3 Materials and Methods

The methodological procedure of the research in this paper consists of several successive steps and is shown graphically in Fig. 1. The first step refers to data collection. In the second step, PCA was applied over the defined set of input variables, which extracted the main components that have the function of input into predictive models. Training and testing of classification models in the IBM SPSS Modeler software environment are performed in the fourth step. Based on the results obtained, the final model is then selected according to the criteria of overall classification accuracy, complexity, and interpretability.

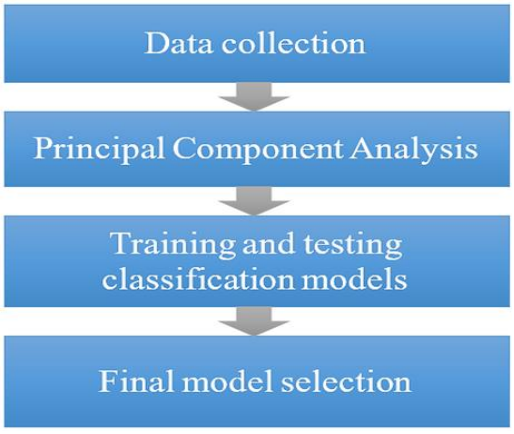

Fig. 1. Methodological procedure of research

3.1 Data Collection

The research is geographically oriented to the area of the section of the M17 main road between the city of Doboj and the Johovac node at the junction of the "9.januar" motorway. This 10 km long section is marked in blue in Fig. 2.

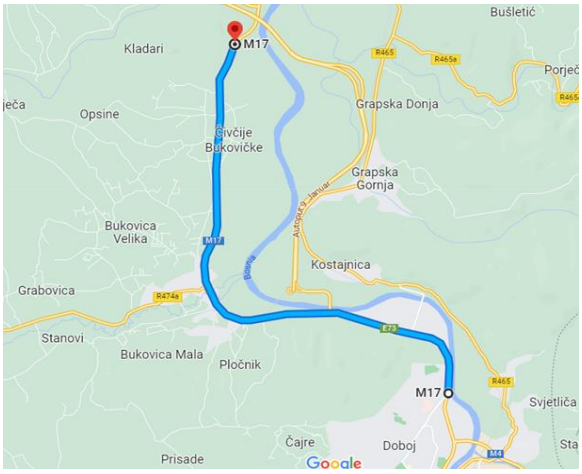

Fig. 2. Geographical area of research [11]

The set of research data was obtained from the mobile operator M:tel Banja Luka based on an officially sent Request (under an official cooperation agreement) which specified the necessary variables [12]. From the obtained database for the 4G LTE network, for the purposes of this research, a total of 1077 values of research variables were extracted for the period between 01.01.2021. and 15.01.2021. with a one-hour sampling frequency [12]. The data are structured into (input/output) vectors. The values of the 17 independent variables given in Table 1 represent the input part, and the value of the dependent variable-delay, the output part of the vector.

Since they are not given in the M:tel database, the continuous delay values are obtained indirectly based on the research presented in [13]. First, the Real Time Video delay value is estimated depending on the number of active users for the exponential rule (EXPRULE) scheduling algorithm (from Figure 6(a) in [13]). As a result of the regression over the estimated values, a cubic function is obtained:

$$y = 8 \cdot 10^{-8}x^3 - 2 \cdot 10^{-5}x^2 + 0.0022x - 0.0063 \quad (1)$$

where y is the delay and x is the average number of users in the cell. This function has a very high coefficient of determination R^2=0.9914 and is shown graphically in Fig. 3.

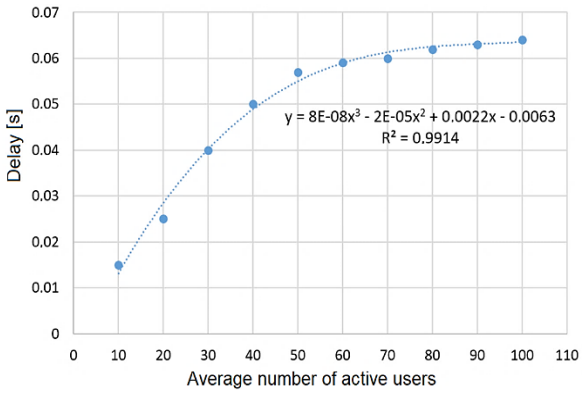

Fig. 3. Approximate cubic delay function

Based on Eq. (1) and the known (from the M:tel database) average number of active users in the cells, delay values were calculated. Classification implies the discrete nature of the dependent variable. Therefore, the values thus obtained were mapped into a discrete scale with three classes according to Packet Delay Budget requirements for individual services [14], which is shown in Table 2.

Table 2. Mapping delay values to classes

Packet Delay Budget	The interval of continuous delay values	Discrete value-class
Real Time Gaming: 50 ms	≤50 ms	1
Conversational Voice: 100 ms	From 50.01 ms to 100 ms	2
Conversational Video-Live Streaming: 150 ms	From 100.01 ms to 150 ms	3

Table 1. Independent/input research variables

Variable label	Meaning
DPUR	Downlink (DL) Physical Resource Block (PRB) Usage rate [%]
ACQI	Average Channel Quality Indicator (CQI)
RRCAtt	Number of attempts by the User Equipment to establish a connection with an eNodeB
CellADT	Cell average DL throughput [kbit/s]
UserADT	User average DL throughput [Mbit/s]
URTR	Uplink (UL) retransmission rate [%]
DRTR	DL retransmission rate [%]
DIBLER	DL Initial Block Error Rate (IBLER) [%]
UIBLER	UL IBLER [%]
CellTVD	Total aggregated DL traffic in the cell [Gbit]
CellTVU	Total aggregated UL traffic in the cell [Gbit]
CellAUT	Cell average UL throughput [kbit/s]
UserAUT	User average UL throughput [Mbit/s]
AUInt	Average UL interference [dBm]
DSCH16QAM_RET	Number of retransmitted TB into DL shared transport channel under 16QAM
DSCH64QAM_RET	Number of retransmitted TB into DL shared transport channel under 64QAM
DSCHQPSK_RET	Number of retransmitted TB into DL shared transport channel under QPSK

3.2 Principal Component Analysis

PCA is one of the most popular techniques used to reduce the dimensionality of large data-sets by linear transformations of variables into a smaller one. They retain most of the information, i.e. the variability of the original set. Thus, the goal is to find or extract new variables - principal components (PCs). They are linear functions of those in the original data set, they are not correlated with each other, and they successively maximize the variance. This comes down to the eigenvalue/eigenvector problem [15]. In this paper, the PCA procedure on 17 input variables is performed in the IBM SPSS Statistics software environment. The main reasons for applying PCA to such a set of variables are as follows: PCA can prevent overfitting; Correlated variables are removed; Machine learning algorithms converge faster; Easy visualization.

3.3 Training and Testing Classification Models

The classification problem in machine learning is solved according to the supervised learning paradigm. The entire set of research data in this paper is divided into two parts: training data and testing data in a ratio of 70:30. This ratio of division was chosen considering that it is one of the most common in research. Structured input-output vectors are processed in the IBM SPSS Modeler software using the Auto Classifier node to automatically create different delay classification models simultaneously. In one pass through the modeling process, Auto Classifier examines different machine learning techniques, with default option settings [3]. In this case, options mean e.g. the number of layers of the neural network, the number of neurons in each of them, the shape and parameters of the classification function, the training algorithm, the stopping criteria, the size of the tree, etc. Supported techniques include neural networks, classification and regression trees (C&R Tree), Quick, Unbiased, Efficient Statistical Tree (QUEST), Chi-square Automatic Interaction Detection (CHAID), C5.0, Logistic regression, decision list, Bayesian networks, Discriminants, Nearest neighbors and Support Vector Machines (SVM). Based on the results of model testing, Auto Classifier offers the most accurate solutions and ranks them according to the overall classification accuracy expressed as a percentage [3].

3.4 Final Model Selection

In addition to the overall accuracy of vector classification, as one of the key indicators of predictive performance, special attention should be paid to its complexity and interpretability when choosing the final model. In general, models with more complex machine learning algorithms are more difficult to interpret. The reason for the complexity of the model, apart from the algorithms, can be the dimensionality of the space of input variables. At the same time, complexity can also be increased by the functions that models need to learn. Although complex predictive models usually provide the best performance, numerous studies give priority to simpler, more interpretable solutions [16,17,18]. Although a formal, mathematical definition of interpretability is lacking in machine learning, the paper [17] lists several definitions of this term. Among them the following stands out as frequently used: "interpretability in machine learning is the degree to which a human can understand the cause of the machine learning model's decision". In recent years, a relatively new field called Interpretable Machine Learning (IML) has emerged, owing to this reason. Methods for transforming black-box machine learning models into white-box are researched in this area [17,19].

In this paper, in addition to the models created over the extracted PC as input variables, classification models are also created over the entire set of 17 variables using the automatic modeling method. In the last step, the prediction performance, complexity, and interpretability of the created models are compared, and the final model is selected based on the previously explained.

4 Results and Discussion

In accordance with the objective and methodology of the research, after collecting the data, it is necessary to reduce the dimensionality of the input space of the machine learning models. Optimization of the number of input variables is achieved by applying PCA over the original set of 17 input variables. As the first result of the PCA, Table 3 shows the value of the Kaiser-Meyer-Olkin (KMO) coefficient and the Bartlett's Test of Sphericity statistics. The sample size is considered adequate if the KMO coefficient is greater than 0.5. In this case, based on the value of KMO=0.795, the same conclusion is drawn: a sample of 1077 input/output vectors is adequate

for PCA application. Bartlett's Test of Sphericity is an indicator of the strength of the relationship between variables. It tests the null hypothesis, which represents the assumption that the correlation matrix is, in fact, an identity matrix in which the diagonal elements are equal to 1, and all others are equal to 0. This test is usually performed before the actual application of one of the data dimensionality reduction techniques. The aim is to confirm statistically that the observed technique can "compress" the data in a meaningful way. It can be seen from Table 3 that the calculated value of the statistic Sig.=0.000 is less than the conventional level of significance α=0.05, based on which it is concluded that the null hypothesis is rejected.

Table 3. KMO and Bartlett's Test of Sphericity

Kaiser-Meyer-Olkin Measure of Sampling Adequacy		0.795
Bartlett's Test of Sphericity	Approx. Chi-Square	27119.791
	df	136
	Sig.	0.000

Fig. 4 shows the values of the standardized data variance (eigenvalue) associated with each extracted component from the initial solution. The most common criterion for choosing the optimal (reduced) set of components is the condition: eigenvalue>1. This means that based on the results given in Fig. 4, it is necessary to take into account the first four significant components (factors) that will represent the entire set of 17 input variables.

Table 4 shows the eigenvalues for the first four components by percentage participation in the total variance of the data. It can be seen that the first component explains 48.47%, the second 13.612%, the third 9.78%, and the fourth component 8.266%. Cumulative, four-component solution explains a total of 80.128% of the data variance.

Table 4. Total variance explained by the four-component solution

Component	Initial Eigenvalues		
	Total	% of Variance	Cumulative %
1	8.240	48.470	48.470
2	2.314	13.612	62.082
3	1.663	9.780	71.862
4	1.405	8.266	80.128

To increase the interpretability of the solution, the components are rotated, and in this case, as most often, an orthogonal Varimax rotation was performed. In addition, Kaiser normalization increased the stability of the solution in the sample. Table 5 shows a rotated matrix of components. The connection of each variable with the component is represented by a weighting factor that can have a value in the interval between -1 and 1. Loading factors less than 0.30 are not shown in the table. Based on these coefficients, a conclusion can be drawn about the method of "grouping" the variables into individual components. This represents the most important result of PCA.

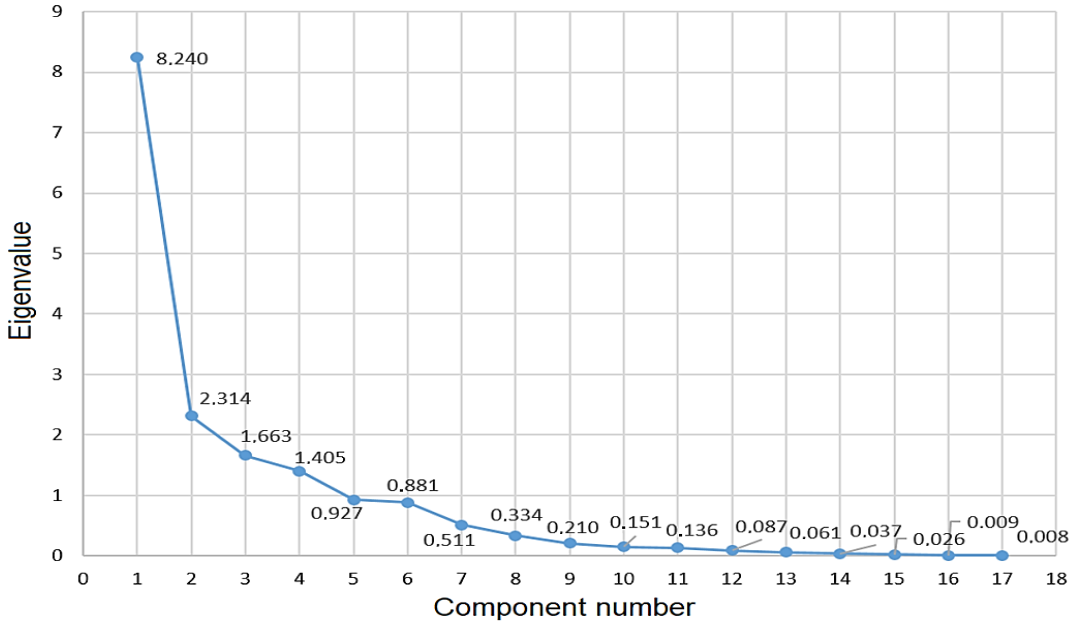

Fig. 4. Values of standardized variance (eigenvalues) associated with individual components

Table 5. Rotated component matrix

Input/Independent variable	Component			
	C1	C2	C3	C4
DPUR	0.956			
DSCH16QAM_RET	0.926			
CellTVD	0.911			
DSCH64QAM_RET	0.837	0.325		0.340
UserADT	-0.800			0.444
DSCHQPSK_RET	0.742		0.413	
CellTVU	0.737	0.561		
RRCAtt	0.597		0.365	0.487
DIBLER	0.580		0.576	
UserAUT		0.843		
CellAUT	0.404	0.834		
AUInt		0.453		
DRTR			0.837	
URTR		0.385	0.729	
UIBLER	0.410	0.474	0.641	
ACQI				0.865
CellADT				0.820

The presented PCA results indicate that it is possible to consider the four extracted components C1, C2, C3, and C4 as inputs to predictive models. They retain information from the original set of input variables to a certain extent. In the IBM SPSS Modeler software environment, data is processed using the Auto Classifier option, to automatically train and test different classification models. The performance for the three best solutions is presented graphically in Fig. 5. It is concluded that Logistic regression with a total classification accuracy of 92.13% is the most accurate model created over the input set of four components obtained by extraction.

Predictive models were also created over the complete set of 17 inputs. Fig. 5 shows that in this case the Neural network model has the highest percentage of correctly classified vectors of 96.759%. The difference between these two models is 4.629% more correctly classified data in favor of the model with 17 input variables. However, due to the many times smaller number of inputs (4.25 times) and easier interpretability, the first model based on the Logistic regression technique is chosen as the final solution. One of the basic goals of the paper was fulfilled in this way because it can be considered that the final model has sufficiently high accuracy. In both cases, the C5.0 model is shown to have the lowest classification performance values.

Logistic regression, as the finally selected model, makes predictions by creating the following set of equations that connect the values of the input fields with the probabilities of association or belonging to each category-class of the output variable:

$$Y_{Class1} = 0\ (the\ reference\ class), \quad (2)$$

$$Y_{Class2} = 4.069 + 4.138 \cdot C1 + 2.273 \cdot C2 + 1.524 \cdot C3 + 1.265 \cdot C4 \quad (3)$$

$$Y_{Class3} = -8.972 + 8.715 \cdot C1 + 4.546 \cdot C2 + 3.793 \cdot C3 + 3.845 \cdot C4 \quad (4)$$

where Y_{Class1}, Y_{Class2} and Y_{Class3} are the class labels of the output variable. Once a model is generated in this way, it is easily used to estimate class membership for new input data. For the values of the independent variables, the probability of belonging to each class is calculated according to the previously mentioned formulas, and the class with the highest probability is assigned as the final value.

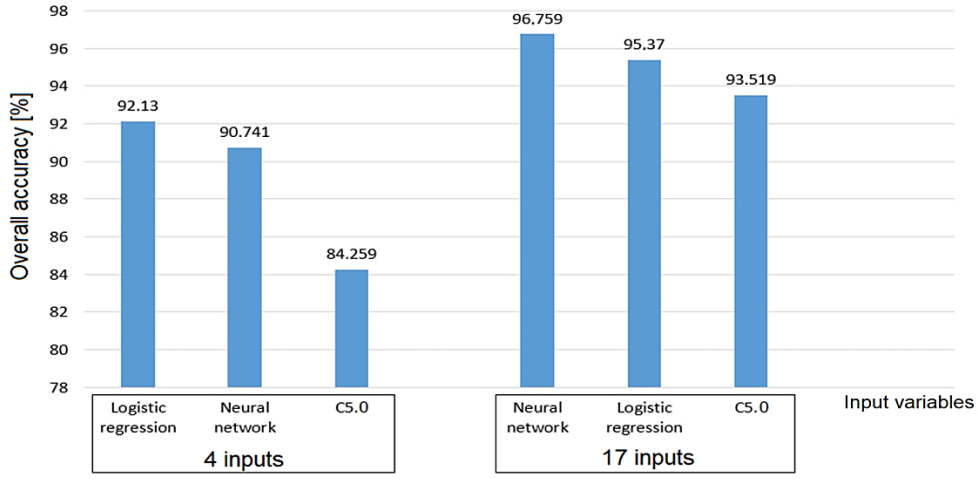

Fig. 5. Classification accuracy values for models created over 4 and over 17 input variables

5 Conclusion

In the paper, the basic objectives of the research that were set were fully achieved. Predictive models were created for the classification of end-to-end delay for real time services in the 4G LTE network. The complete available set of 17 input variables was reduced to four PCA extracted variables. The machine learning model was trained and tested by the supervised learning paradigm on the complete available set of inputs, as well as on the extracted components. In the last step, based on the defined methodology and research objectives, the final model was chosen according to the criteria of overall accuracy, complexity and interpretability. For the full set of input variables, the best prediction performance was shown by the neural network model with 96.758% correctly classified vectors from the test set. The results of testing the classification models for four inputs showed a slightly lower percentage of classification accuracy. As the best model, in this approach, Logistic Regression has a classification accuracy of 92.13%. These model performances in both cases, according to the research objectives, are considered sufficiently high and acceptable. When choosing the final solution, in addition to the specified classification accuracy, the complexity and interpretability of the model was taken into account. In this paper, the Logistic Regression model with four inputs was chosen as the final model because it is less complex and more interpretable than the neural network model. The feature extraction technique for reducing the dimensionality of the space of input variables has shown its effectiveness. In addition to interpretability, the final model retained a large percentage of the information from the initial data set. The four extracted components explain a total of 80.128% of the data variance. These components indirectly through delay and measures of rigidity R1, R2, R3, R4 and R5 affect QoE and sustainable urban mobility. Delay as an important indicator of network performance is related to and to a certain extent affects SUMP performance such as quality of life, urban environment, effective use of limited resources, etc. Therefore, knowing the value of the delay in the future using predictive models also means knowing the degree of SUMP fulfillment in a certain segment. Greater accuracy of delay classification provides greater accuracy of SUMP fulfillment degree prediction. Future research directions can be oriented towards the application of other individual Feature Extraction techniques, and the comparison of the obtained results.

References

[1] N. K. Madi, & M. Madi, "Analysis of downlink scheduling to bridge between delay and throughput in LTE networks", In *2020 7th international conference on electrical and electronics engineering (ICEEE)*, IEEE, April 2020 (pp. 243-247).

[2] J. P. Singh, P. Dutta & A. Pal, "Delay prediction in mobile ad hoc network using artificial neural network", *Procedia Technology*, *4*, 201-206, 2012, https://doi.org/10.1016/j.protcy.2012.05.030

[3] M. K. Banjanin, M. Stojčić, D. Danilović, Z. Ćurguz, M. Vasiljević, & G. Puzić, "Classification and Prediction of Sustainable Quality of Experience of Telecommunication Service Users Using Machine Learning Models", *Sustainability*, *14*(24), 2022, 17053.

[4] I. de-la-Bandera, D. Palacios, J. Mendoza & R. Barco, "Feature extraction for dimensionality reduction in cellular networks performance analysis", *Sensors*, *20* (23), 6944, 2020. https:// doi.org/10.3390/s20236944

[5] A. H. Ahmed, S. Hicks, M. A. Riegler & A. Elmokashfi, "Predicting High Delays in Mobile Broadband Networks", IEEE Access, 9, 168999-169013, 2021.

[6] A. S. Khatouni, F. Soro & D. Giordano, "A machine learning application for latency prediction in operational 4g networks", In *2019 IFIP/IEEE Symposium on Integrated Network and Service Management (IM)*, April 2019, (pp. 71-74). IEEE.

[7] W. Zhang, M. Feng, M. Krunz & H. Volos, "Latency prediction for delay-sensitive v2x applications in mobile cloud/edge computing systems", *In GLOBECOM 2020-2020 IEEE Global Communications Conference*, IEEE, December 2020, (pp. 1-6).

[8] J. Brown & J. Y. Khan, "A predictive resource allocation algorithm in the LTE uplink for event based M2M applications", *IEEE Transactions on Mobile Computing*, *14*(12), 2015, 2433-2446.

[9] W. K. Lai, C. W. Hsu, T. H. Kuo & M. T. Lin, "A LTE downlink scheduling mechanism with the prediction of packet delay", *In 2015 Seventh International*

Conference on Ubiquitous and Future Networks IEEE, 2015, July (pp. 257-262).

[10] R. Zhohov, D. Minovski, P. Johansson & K. Andersson, "Real-time performance evaluation of LTE for IIoT", *In 2018 IEEE 43rd Conference on Local Computer Networks (LCN)*, 2018, October (pp. 623-631), IEEE.

[11] https://www.google.com/maps/dir/44.7438539,18.0953169/44.8027333,18.048289/@44.782805,18.0279954,12.26z/data=!4m2!4m1!3e0

[12] M. K. Banjanin, M. Stojčić, D. Drajić, Z. Ćurguz, Z. Milanović, & A. Stjepanović, A., "Adaptive Modeling of Prediction of Telecommunications Network Throughput Performances in the Domain of Motorway Coverage", *Applied Sciences*, *11*(8), 3559, 2021. https://doi.org/10.3390/app11083559

[13] J. I. A. Y. Yaqoob, W. L. Pang, S. K. Wong, & K. Y. Chan, "Enhanced exponential rule scheduling algorithm for real-time traffic in LTE network", *International Journal of Electrical and Computer Engineering (IJECE)*, *10*(2), 2020, 1993-2002.

[14] ETSI TS 123 203 V9.7.0 (2011-01), "Technical Specification Digital cellular telecommunications system (Phase 2+); Universal Mobile Telecommunications System (UMTS); LTE; Policy and charging control architecture (3GPP TS 23.203 version 9.7.0 Release 9)" https://www.etsi.org/deliver/etsi_ts/123200_123299/123203/09.07.00_60/ts_123203v090700p.pdf

[15] I. T. Jolliffe & J. Cadima, "Principal component analysis: a review and recent developments", *Philosophical transactions of the royal society A: Mathematical, Physical and Engineering Sciences*, *374*(2065), 2016, 20150202.

[16] M. A. Ahmad, C. Eckert, A. Teredesai, "Interpretable machine learning in healthcare", *In Proceedings of the 2018 ACM International Conference on Bioinformatics, Computational Biology, and Health Informatics*, Washington, DC, USA, 29 August–1 September 2018, pp. 559–560.

[17] T. A. Abdullah, M. S. M. Zahid & W. Ali, "A review of interpretable ML in healthcare: taxonomy, applications, challenges, and future directions". *Symmetry*, *13*(12), 2021, 2439.

[18] B. Dherin, M. Munn, M. Rosca & D. Barrett, "Why neural networks find simple solutions: the many regularizers of geometric complexity", *Advances in Neural Information Processing Systems*, *35*, 2022, 2333-2349.

[19] G. Stiglic, P. Kocbek, N. Fijacko, M. Zitnik, K. Verbert, L. Cilar, "Interpretability of machine learning-based prediction models in healthcare", *Wiley Interdiscip.-Rev.-Data Min. Knowl. Discov.*, 2020, 10, e1379. doi:10.1002/widm.1379.