MACHINE LEARNING MODELS IN THE CLASSIFICATION AND EVALUATION OF TRAFFIC INSPECTION JOBS IN ROAD TRAFFIC AND TRANSPORT

PhD Milorad K. Banjanin[1,2]; MSc Aleksandar Đukić[3]; PhD Mirko Stojčić[3]; MSc Milan Vasiljević[2]
[1] University of Novi Sad, Faculty of Technical Sciences Novi Sad, Serbia, milorad.banjanin@ff.ues.rs.ba
[2] University of East Sarajevo, Faculty of Philosophy Pale, Bosnia and Herzegovina, milanvasiljevic84@gmail.com
[3] University of East Sarajevo, Faculty of transport and traffic engineering Doboj, Bosnia and Herzegovina,
mirko.stojcic@sf.ues.rs.ba, a.djukic@inspektorat.vladars.net

Abstract: *Traffic inspection in the road transport of people, products, and values is the central construct in the subject of this research. It is discursively designated as a specialized function of a certain traffic system and organizationally situated as a unit of the state or republican inspectorate for inspection affairs, responsible for the safe and reliable functioning of traffic and transport entities in a certain geo-area. Inspection jobs are predominantly oriented to the supervision and control of the exploitation of infrastructural facilities and dynamic facilities with a special focus on the variables of traffic flows and transport processes and people as participants in traffic with different roles. The machine learning models created in this paper learn the conceptual, constructional, visual, and relational elements of the form of physical (stationary and dynamic) objects and their attributes, in the field of interaction with the biological, software, and cyber-physical components of the traffic system to which the tasks of traffic inspection are related. In addition to people, the roles in traffic inspection work are also realized by operational technologies as software and cyber-physical agents that function synergistically in unstable situational contexts of traffic reality. The aim of the research is to solve well-posed problems of classification and evaluation of traffic inspection jobs in the function of greater traffic safety and more efficient transport. The hypothetical setting of the research is that machine learning techniques (Artificial Neural Networks (ANN), Decision Trees (DT), k-Nearest Neighbors (k-NN), Support Vector Machines (SVM)) can successfully classify and evaluate traffic inspection jobs on platforms for their digitization and digital treatment of research data. Data for digital processing is collected from very different information sources that are located in the space of the road belt of traffic roads and other objects with which they interact..*

Key words: *traffic inspection, traffic flows, transport processes, machine learning techniques and models, classification and evaluation of traffic inspection jobs, operational technologies, digitization of jobs, digital processing of research data*

1. INTRODUCTION

Traffic and transport are, above all, economic activities of general importance for the individual, collective, social, local-communal, state and global level of functioning of individual countries. They are structured as complex systems that are designed, built and function in a dynamic density of relationships, relationships and content of people's lives and work with the aim of producing and providing traffic and transport services. Logically and structurally, traffic and transport systems were created as engineering systems with the inevitable influence of social, legal-regulatory, innovative-technological, contextual-relational and other environmental factors. The constructional modular structure of the system means that individual modules can function and be studied as independent units that are connected to other modules in different ways. One of the modules of that system, as a specialized supervisory-control function, is the traffic inspection (SI), whose roles of supervision and control are positioned in the interactive dimensions of the content and contextual structure of specialized jobs. According to [1] the main job categories in the **content structure** are: knowledge (conceptual, descriptive, procedural, functional at the level of data, information, functional acquired knowledge, applied knowledge, intelligence, wisdom and capacity for cooperation and sharing in learning episodes and practice episodes through multi-agent collaborative interactions); skills (conversational, correspondence, visual, manipulative, digital and multimodal in multi-agent network communication); work and work processes (actions, activities, operations, events, transactions and transactional events) and technologies (operational, manipulative, information, communication, software, network applications and embedded systems). In the contextual structure of SI jobs, the main categories are: orientation to (content, task, target, activation and transfer function, location, infrastructure objects, mobile objects), to strategies and traffic culture (Figure 1).

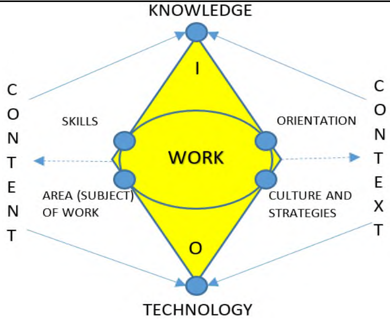

Figure 1: The main components of the job in the content and contextual structure [1]

SI tasks are categorized as simple, complicated (complex) and complex and are performed by solving tasks in identically categorized situational contexts with singular or plural roles of human agents (HA), software (SA) and cyber-physical agents (CPA) on a multiagent communication (MAC) platform in the defined traffic geo-space. In this paper, the geo-space of the road traffic and transport system in the Republic of Srpska (RS) and BiH was chosen for the case study for which machine learning (ML) models are created for the classification and evaluation of monitoring and control tasks. The system and process structuring and functioning of traffic and transport in that area takes place according to technical-technological standards, norms, algorithms and protocols, and in the organizational-management and social-contextual domains on the basis of legal norms and rules prescribed by the set of laws of the RS and international norms, declarations, regulations and directives of the European Union (EU). The hypothetical setting of this research is: "The roles of new knowledge, skills and technologies in information processing where perceptive and intuitive processing that is based on procedural, reflexive and communal or contextual learning dominate are insufficiently understood by the employees in SI institutions. In work processes and SI technologies, multi-aspect cognitive and formative explicit processing of the components of solutions to business problems and situations through integrative, generative and synergistic learning in network MAC is necessary. In multimodal interactions MAC multifaceted knowledge, skills, technologies and processes in HA roles are expanded and perfected through learning episodes and practice episodes and connected with the techniques of machine learning methods SA and CPA in induced experiences over a set of problem solving examples that are memorized in the KB- knowledge base SI" (Figure 2.)

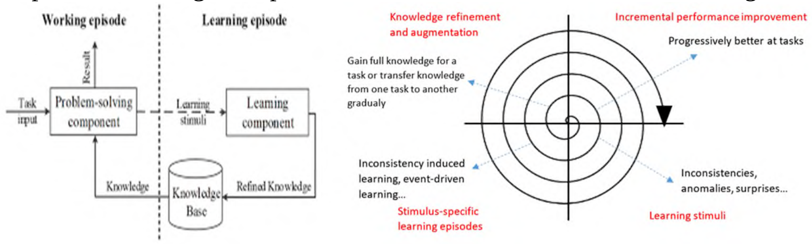

Figure 2: A conceptual model of KB SI induced by agents in learning episode, practice episode and ML models [2]

Based on this hypothetical setting, the research is oriented towards identifying and connecting interactive situational factors as units from certain paired empirical paradigms in the field of traffic, misdemeanor and criminal law, construction, electric power, traffic, transport and communication infrastructure, spatial information infrastructure, vehicles and their modalities, human factor in singular and plural format, hardware, software and network technologies and environment in network MAC. The purpose of connecting data from the units of the mentioned paradigms is to model optimal solutions to the problems of preventive, action and educational activities of SI. The final goal is that SI, with a competitive business contribution, dynamizes the strengthening of the road traffic safety system in the RS in cooperation with other partners to solve problems and tasks of inspection supervision and control through proactive and permanent engagement of agents in learning episodes, practice episodes and machine learning models. According to [3], one of the challenges in machine learning is developing systems that never stop learning new tasks and continuously improve their performance [3]. In a learning episode, the learning component uses stimulus-specific algorithms or heuristics to augment or revise existing knowledge, resulting in improved performance P [2]. After the agent continuously improves experience (E), it can satisfactorily perform tasks [2]. In the practice episode, at the end of the task, the agent examines the appropriate knowledge in the knowledge base (KB), i.e. the experience E based on which the problem-solving component generates the result for the task [2,3]. By evaluating the results of the processing using the performance metric P, it is possible to evaluate "how well the agent works", with the fact that success in solving the task depends on the quality of the experience and stability, i.e. the variability of the environment [2,3].

2. MATERIAL AND METHODOLOGY

The problem model that the agent analyzes and learns how to solve consists of an agent, a state, and a set of actions related to a particular state. By performing an action, the learning agent can move from one state to another, and the transition is provided with a certain "reward". The goal of the agent is to maximize the "total reward" by learning which action is optimal for each state included in the problem model. In this paper, models of problems in SI jobs are coded into three main categories - simple, complicated and complex, and are performed in tasks that are identically categorized in situational contexts. Coding and creation of code names of selected and evaluated linked multidimensional data in this research was carried out using two-dimensional tables (Figure 3.) on the basis of which the collected data on the institutional evaluation of quality parameters (WHAT?, WHO?, WERE?, HOW?, WHEN?, WHY?) of the results achieved by individual agents in the personnel composition of SI who solve professional problems and business tasks.

Code	**Code name of the job/assignment** (WHAT?)	(WHO?) **INSPECTORS**					(HOW?) **ASPECTS OF KNOWLEDGE**						
	{supervises (**n**), encodes (**k**), processes (**p**), solves (**r**), selects (**s**), evaluates (**e**), authorizes (**a**)} →	**GR**	**R**	**KR**	**G**	**O**	**p**	**i**	**z**	**PZ**	**I**	**M**	**C**
001													
002													
...													
250													

Figure 3: Multidimensional connected data in traffic inspection jobs-tasks

Tasks are defined with attributes of structured data for digital processing in the operational roles of monitoring (n), coding (k), processing (p), analytical solution (r), selection (s), evaluation (e) and authorization (a). According to the content of the task, the implementation of which solves the problem in the domain of competence of the SI and the target function, supervision can be legal, administrative, technical, technological, functional, process, project, investment, contractor, construction-traffic, and according to the scope of execution, individual, group or team and multi-agent. Accordingly, the authorized performers of transparent SI roles are human agents-inspectors in the positions of chief republican inspector (GR), republican inspector (RI), inspector-coordinator of the region (KR), city (GI) and municipal inspector (OI) who function at the level of individual aspects of knowledge that structure the cognitive continuum.

Aspects of knowledge in the cognitive continuum are data (p), information (i), functional knowledge (z), applied knowledge (PZ), intelligence (I), wisdom (M) and capacity (C) to share knowledge in a multi-agent team, group or ensemble, in MCA with objects, processes and other factors using multimodal interactions. In multimodal communication, SI uses conversation modes (face-to-face (F2F), telephone and viber-network conversations), e-correspondence (e-mail, SMS, MMS), analog and digital data transmission, presentation, sharing and exchange of written documents, diagrams, graphs, tables, photo-documentation, video and audio records and streaming, records from certain sensors, cameras, radars, direct observations, mapping of data and functions, etc [4]. Certain forms of cooperation (cooperation, coordination, collaboration) with business-technical and other partners take place in a time frequency of daily, weekly, quarterly, semi-annually, annually or according to special criteria of categorization and selection of related data in SI affairs. Therefore, an important parameter for all modes, contents of tasks and problems and contexts of situations is their repeatability.

In SI's interactions with agencies, companies, entrepreneurs, business alliances, chambers of commerce and entrepreneurship, linked structured data in the form of queries is used, demands, orders, beliefs, official notes, correspondence rules, reports, decisions and other documents as part of functional and non-functional specifications for a certain software system in the SI information system. The main content or dispositif in the mentioned documents is built by symbols, words and sentences as premises that represent transitions between the states of the system. Symbols, words and sentences must meet certain criteria of *dynamic semantics and practical inference* [5]. The basic criterion is validity, which implies a quality of meaning based on truth or reason, or a conclusion that can be accepted. In particular, data validity, as one of the six dimensions of data quality, "simply means how well the data meet certain criteria, which often evolve from analysis of previous data as relationships and problems are discovered." According to [5,6] "a practical conclusion is a conclusion whose at least one premise is a sentence that sets the goal". For example, the goal of analytically solving the problem of the practical conclusion of the supervision of some object of the traffic infrastructure is to check the validity of the supervision of operational readiness, functional reliability and reliability of the functioning of the supervised object or component of the problem. The term dynamic semantics indicates that there are many types of sentences and countless different ways of using what are called «symbols», «words», «sentences». Wittgenstein [7] points out that, although this multiplicity is not something unchanging and given once and for all, but new types of language, new language games, "the importance of dynamic semantics is that it offers a way of thinking about the meaning of sentences in natural language, i.e. that different types of semantic actions belong to different modes". The number of imaginable semantic actions exceeds the number belonging to the concept of the basic modes of the statement, question and command type. This means that "the standard tripartite division of modes into indicatives, interrogatives and imperatives can be and perhaps should be retained" [5,6] with reliance on "intra-modal and inter-modal relations of sentences".

The mentioned modes are frequently used in simple tasks of SI, i.e. tasks that solve problems in simple situations for which one path leading to the solution is necessary and sufficient. On the other hand, the existence of multiple paths represents a fundamental difference between simple and complex jobs and tasks. It is generally considered that complex jobs are defined by situations that meet the following conditions: a complete hierarchical structure of all the paths used to solve the tasks in the job can be made; solutions in practice may be difficult to implement even though they can be defined in the job design; the hierarchical structure of work usually becomes a network of situations with multiple and multidirectional connections between the paths used to solve tasks; multidirectional connections within the network, with multiple paths from the start point to the end point make the job more complex because a complex job for each task may have several paths leading to the solution. As an example of a product of complex work in traffic inspection, the software of the Tax Administration handles every irregular behavior of SI agents. The problem is not defining work situations or avoiding them, but generating a path to solve the problem.

Complex jobs contain too many factors, so they cannot be fully analyzed, so it is impossible to provide a complete set of information or accurately define paths through individual job situations. This means that complex work situations contain a lot of ambiguities and small but important differences in the available or required information. This fact, if nothing else, requires the presence of HA in the process, especially when computer software simply cannot handle unpredictable situations and provide reliable answers to ambiguities. A complex job differs from a simple one in that it is constantly under the influence of various external factors on the output variables of the inspection process. External factors add a dose of unpredictability that greatly complicates the analysis. Cilliers (1998) described a complex job involving different situations with the words: "the interaction between the system's constituent parts and the interaction between the system and its environment are such that the system as a whole cannot be fully understood through component analysis. Moreover, these relations are not immutable, but shift and change" [7].

The difference between simple and complex jobs on the example of evaluation of inspection reports can be determined by comparing the job of an agent, who enters data and an agent of a business analyst. The data entry agent does a simple job, as forms appear on the screen and it enters data into certain system windows. Each value goes to a certain place provided by the algorithmic structure. Problematic forms and incomplete information require external intervention (refill the form or let the agent-supervisor decide what to do) as they cross the established boundaries of a simple situation. On the other hand, the business analyst, who receives reports generated from such data, works in a complex situation. He is tasked with using that information to determine how certain segments of the business are performing and to suggest necessary improvements. It is a well-known principle that the business analyst must use more information in the analysis than the volume of data entered in digital processing. Any external information that cannot be obtained through data entry, and which is necessary for interpretation, can prove to be more influential than the data.

3. MULTIAGENT LEARNING IN TRAFFIC INSPECTION JOBS

3.1. Learning of human agents and the development of the cognitive continuum of knowledge in learning episodes and practice episodes

From the organizational aspect, the Traffic Inspection can be seen as an institution for the integration of knowledge or as a dynamic system, with knowledge-based activities, which deals with organizational knowledge and the ability to generate new organizational knowledge ». According to [8] Dorothy Leonard-Barton (2004), in her analysis of knowledge in the organization, she distinguishes the knowledge of employees and knowledge in physical systems (which is embedded in machines, databases). Also, it makes a distinction between employee knowledge that is acquired in learning episodes and employee skills that they possess (knowledge manipulation skills) and develop new ones in practice episodes. Employee knowledge is knowledge that one person or a set of persons (eg, a group, team, or other social entity) provides when applying (using) knowledge in the organization, or that becomes available when performing activities of applying and improving knowledge in the organization. Newer research indicates the importance of understanding different aspects or types of knowledge of individuals in the cognitive continuum and the level of consciousness in their behavior and actions, the performance focus of the activity and the time perspective that corresponds to a certain type of learning and the achieved level of consciousness. Different levels of consciousness develop correspondingly to the types of learning of certain levels or aspects of knowledge. Knowledge of employed individuals, organizational and social knowledge "weave a complex network of knowledge".

The basic way of organizing the complex network of human knowledge at the level of the individual is designated as his «cognitive continuum». According to [9], "the cognitive continuum is formed by learning simple, complicated and complex content, solving simple, complicated and complex tasks in work, so that the structure of knowledge is not hierarchical, but simply a framework for understanding the network of human knowledge, which does not know about hierarchy and the effects of learning and the range of certain aspects in the direction of development of the continuum, are interpolated on the line of learning". Therefore, certain types or aspects of knowledge in the cognitive continuum are: p, i, z, pz, I, M, C and these are not elements of a hierarchical structure but interdependent elements of an individual network structure of all aspects of knowledge. Each aspect or type of knowledge or cognition is supported by a corresponding learning activity, respectively: instinctive, procedural, reflexive, communal or contextual, integrative, generative and synergistic learning. Certain aspects of knowledge are determined with an action-performance focus, respectively: data collection, execution of procedures (procedurality-algorithms, strategies in activities, action plans, programs and methods), knowledge for functioning (functionality), management, integration, regeneration - renewal and cooperation (cooperative, coordinating, collaborative).

Sight - an aspect of knowledge	Learning type	Performance focus	Time perspective	Consciousness -level
DATA	Instinctively	Data collection-feedback	Current	Feeling consciousness
INFORMATION	Procedural	Procedurality-efficiency	Very short term	Receptor-sensory
KNOWLEDGE	Reflexively	Functionality-efficiency	Short term	Consciously reflexive
APPLIED KNOWLEDGE	Systemic-structural	Productivity-management	Medium term	Communal-contextual
INTELLIGENCE	Integrative	Optimal integration	Long term	Structural
WISDOM	Generative-open interactional	Restoration-integrity of connections	Very long term	Ethical
Knowledge sharing capacity of the knowledge and learning community	Synergistically	Cooperation in the form of cooperation, coordination, collaboration	Timeless	Universal

Figure 4: Tabular presentation of aspects of knowledge of human agents in the cognitive continuum [9]

The effects of learning in SI jobs can be demonstrated with improved performance, i.e. performance goals are respectively marked as: feedback on information gathering, work efficiency with an emphasis on developing, monitoring and completing tasks, effectiveness of functions with an emphasis on effective actions and elimination of inconsistencies, productivity in the application of knowledge and understanding of multiple variable effects, optimization of activities that are viewed as a whole, and adaptability to changes in the environment, integrity of values, visions and missions with an understanding of purpose and cooperation to achieve performance with a consistent understanding of values in a wider context. The learning activity for each aspect of knowledge, with a corresponding action focus and performance goal, is dimensioned with a time horizon, respectively: immediate-current, very short, short-term, medium-term, long-term, very long-term and intergenerational or timeless horizon, which also follows the development consciousness at the level of: feelings, physical sensitivity, self-reflection, communal, structural, ethical and universal consciousness (See table in Figure 4.)
The analyzed human cognitive continuum of knowledge, its various aspects with corresponding types of learning, performance focus, development of the level of consciousness and time perspective of the learning process, and MAC constitute the integrity of the environment and intelligent systems whose behavior is investigated. As can be concluded, the outline overview in tabular form visually describes certain aspects of the knowledge complexity structure with appropriate characteristics and performance focus [9]. From a business perspective, this means that certain types or aspects of the knowledge continuum (data, information, knowledge, applied knowledge, intelligence, wisdom and the capacity to share knowledge and learning) are not separate, isolated, discrete (interruptible) levels of knowledge, but are determined by situational positions of cognitive capacity of individuals. It is clear that the environment is much more than the physical world consisting of individual objects on the street, lighting conditions, architectural diversity, and that its content is also filled by users of products and services who have their own identification, locations, goals and activities, as well as the social environment in which they traditionally analyze the parameters of conventional space with reactive, proactive and co-active interactions of humans with other systems (infrastructural facilities, roads, cars, networks...) or other objects. Environmental development refers to automatically changing the situation and context. This primarily refers to the "appropriation" of entities, roles and relationships on the basis of which simple, complicated and complex situations and contexts are created with the new interaction paradigm of Ambient Intelligence (AmI). AmI means an intelligent space in which people can conduct multiple - multimodal dialogues with objects that have some of their characteristics [4]. Interactions, as ways in which objects communicate with the aim of achieving expected behavior and performing a specific task, can be represented by specifying the messages that they forward or exchange with each other. In this communication, the objects act as an intelligent environment whose conversational interface is combined with a real user interface that uses the situational context.

3.2. Machine learning models for solving tasks of inspection supervision and control

In this work, selected ML techniques are used on the multiagent communication platform, namely [5]:
-*Artificial neural networks (ANN)* model the functioning of the human brain. through basic process elements - artificial neurons that are interconnected in a layered architecture. The input layer has the function of accepting the values of the input variables and forwarding them to a certain number of neurons in the hidden layer, and each neuron of this layer forwards its output to the inputs of the neurons in the next layer and so on to the output layer. The output layer, as the name suggests, provides the final output from the network. Hidden layers perform data processing by assigning weighting factor values to all inputs, which are modified during the training process.
-*Decision Trees* simulate the way people make decisions. In addition to regression, they can solve classification problems if the dependent variable is categorical, so their other name is classification and regression trees (Classification And Regression Trees - CART). During the training process, starting from the root of the tree, nodes are further branched to child nodes, and the target value of the dependent variable is represented by a terminal node - a leaf (Terminal Node). Therefore, by training the model, the conditions represented by the branches are tested, and based on this, the original data set is split into subsets, resulting in a tree. On each derived subset, the process is repeated recursively until the subset of a particular node has all the same values of the output variable, or when further branching no longer contributes to improving the result.
-*Support Vector Machines (SVM)* is a classification technique in which the initial problem is transformed into a search for hyperplanes that separate two classes using a kernel, which defines the interspace. In this new space, it is assumed

that the problem to be solved is linear [10]. SVM maps the training examples to points in space, with the goal being to maximize the space between two classes. New examples are mapped into that space and classification is performed depending on the side of the kernel where the point is found.

- *k-Nearest Neighbors (k-NN)* technique can be used for regression or classification problems, and is most often used as a classification algorithm. His work is based on the assumption that similar points can be found close to each other. In other words, k-NN classifies the new data into the class that is most similar to the available categories [10,11].

4. RESEARCH RESULTS AND DISCUSSIONS

Given that one of the key steps for defining a model or an ensemble of ML models is the selection of independent/input and dependent/output variables, Table 1 gives a systematic presentation of the variables by individual models for solving IS monitoring and control tasks.

Table 1: Input and output variables of the model for solving SI monitoring and control tasks [12]

Supervision and control tasks	**Input variables**	**Output variable/variables**
1. Consistent application of traffic regulations	Type of regulation; Geographical location; Subject of control.	Number of offences
2. Representation of technical standards in traffic infrastructure facilities	Municipality/city; Control object: 1)Physically stable and some mobile objects of traffic and transport; 2)Legal-regulatory, energy-economic, security-protective and ecologically-sustainable facilities; Organization in charge of maintenance; Age of the object; Period of the year.	Safety assessment of facilities; Control result.
3. Traffic-safe exploitation of dynamic facilities in public transport of people by buses, taxis and vehicles of rent-a-car agencies	Number of regular controls of technical correctness; The number of extraordinary inspections of technical correctness; Prices of services; Regularity of departures in public transport;	Level of satisfaction of users of transport services [13]
4. Intelligent and legal behavior of human agents in singular and plural roles of traffic inspection	Level of government; Competences; Role in the team; Location where it operates; Level of functional knowledge.	Results of the inspector's internal control; Agent grades.
5. Functional implementation of operational software and cyber-physical technologies in the context of traffic flow security	Annual investments; The level of use of conventional business technologies; The level of use of Information Technologies (IT); Level of use of communication technologies (CT); Use of personal networks; The level of use of information and communication technologies (ICT); The level of use of software applications in CPA interactions.	Achieved results in terms of safety; Number of traffic accidents.

The main emphasis in this research is on the fourth task shown in Table 2, which refers to Intelligent and legal behavior of human agents in singular and plural roles of traffic inspection. The table defined independent variables in the machine learning model for predicting HA scores are:

1. Level of government (Level), which can be municipal or republican;
2. Competence, expressed through attributes such as: experienced, communicative, meticulous, strict...;
3. Team role, which can be: republican inspector, chief republican inspector, regional coordinator, city inspector or municipal inspector;
4. The location where it operates (Location), which means the city/municipality in the RS;
5. Level of functional knowledge, which can be: data, information, knowledge, applied knowledge, intelligence, wisdom, capacity for knowledge sharing and exchange in complex situations.

The dependent or output variable is represented by the agents' discrete ratings, which can have an integer value in the interval between 1 (unsatisfactory) and 5 (excellent). The main goal is to create predictive models that associate certain values of input variables (input vector) with HA ratings, i.e. classify the input vectors. The training of classification models is performed according to the supervised learning paradigm, where the entire data set is structured into 48 input-output vectors and divided into two parts: a) training data, which make up 90% of the total set and b) data for testing the performance of the model, which make up the remaining 10% of the available set.

Classification predictive models are created in the IBM SPSS Modeler software environment using the Auto Classifier option. This option enables automatic creation and testing of various machine learning techniques, with default parameters. In the last step, the software ranks the tested solutions according to the criteria of correlation and overall classification accuracy. 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, k-Nearest Neighbors (k-NN) and Support vector machine (SVM). The results of training and testing different models using the Auto Classifier procedure are shown in Table 2. Models C5, k-NN, SVM, CHAID and C&R Tree have the highest and equal value of total classification accuracy, which is 88.889%. The performance of the neural network based model is zero. When choosing the final solution, in addition to the overall accuracy of the classification, it is necessary to take into account the complexity expressed by the number of input variables or predictors on which the model was created. Considering that, as a final solution, according to Table 2, an interpretable C5 model with only one predictor is chosen. Figure 5.a shows the structure of the selected C5 model, which is based on the ML Decision tree technique and performs prediction by dividing the sample based on the fields that provide the maximum value of information gain. The division is done by branches where the initial branch or root includes all the data. Further, the root is divided into children branches based on the set of rules given in Figure 5.b.

Table 2: Results of the Auto Classifier procedure

Model	C5	k-NN	SVM	CHAID	C&R Tree	Logistic regression	Bayesian network	Neural network
Overall Accuracy [%]	88.889	88.889	88.889	88.889	88.889	11.11	11.11	0
Number of predictors	1	5	5	1	2	5	5	5

On the left side of the rule, the conditions and restrictions related to the values of the input variables are defined, and on the right side there are classes-grades to which the input values belong. As can be concluded from Figure 5.a and Figure 5.b, Level of functional knowledge is the only significant input variable that figures in the input space of the C5 model. Each sub-branch can be further divided into sub-branches etc. At the lowest level are terminal branches (or leaves) that cannot be divided further. Each node with which the branch ends displays statistics in the form of the percentage of classified vectors [14].

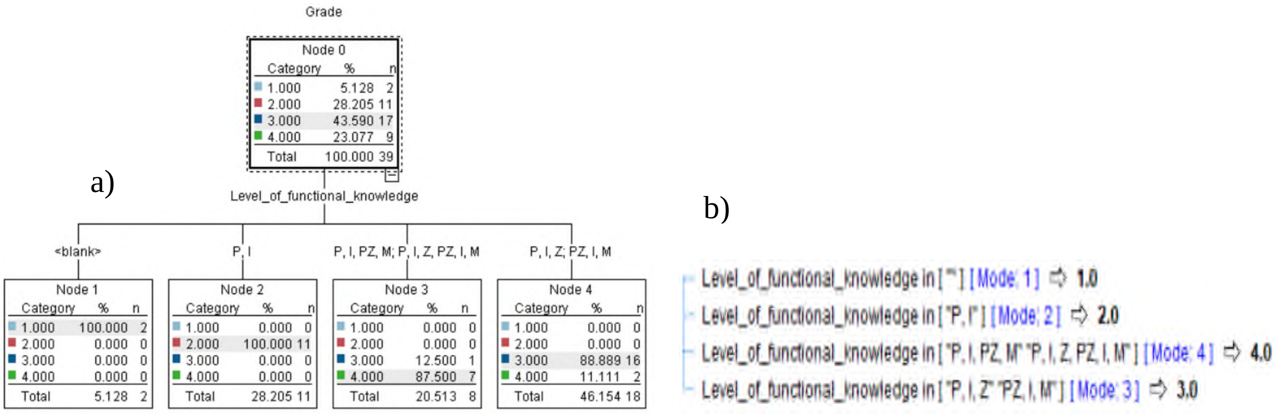

Figure 5: Selected C5 model: **a)** C5 model tree structure; **b)** Set of rules for branching the C5 model tree (where p, i, z, pz, I, M are connected in the nodes of the decision tree)

Figure 6 graphically shows the test results of the selected C5 model. In the second point represented on the x-axis, ie. for the second input vector from the set of 9 input-output vectors for testing, the model gave an incorrect prediction of the HA score. Instead of 4, the input data is associated with a rating of 3 (which can be considered a partial confirmation of the research hypothesis) while in all other cases C5 gave a correct prediction, which corresponds to an accuracy of 88.889%.

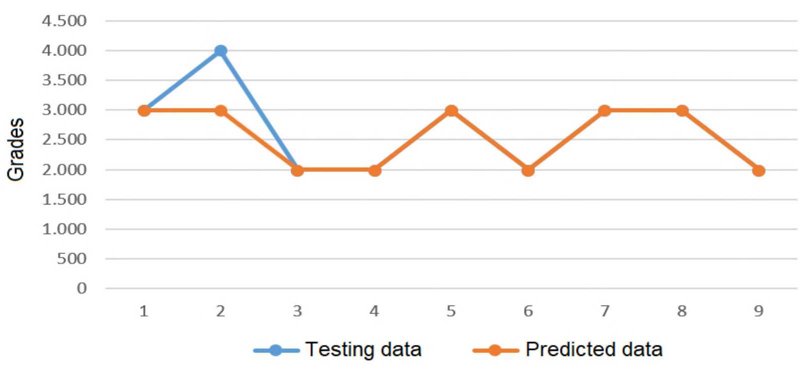

Figure 6: Data set testing results in the selected C5 model

5. CONCLUSION

In the work, special attention is paid to the modeling of Intelligent and legal behavior of human agents in singular and plural roles of traffic inspection. Using the automatic modeling method in the IBM SPSS Modeler software, several models were trained and tested for the classification of input data according to HA scores. Their ranking was performed according to the overall classification accuracy, and the C5 model based on Decision Trees was chosen as the final solution. In addition to the highest percentage of correctly classified vectors of 88.889%, this model was created with only one significant input variable - Level of functional knowledge, which represents one of the advantages of this model in terms of complexity and interpretability. Future research can be oriented towards modeling other tasks of traffic inspection supervision and control using machine learning models.

REFERENCES

[1] M. Banjanin, P. Gojković, „Analitičke procedure u inženjerskim disciplinama“, Univerzitet u Istočnom Sarajevu-Saobraćajni fakultet, Doboj, 2009, UDK COBISS.BH-ID 781336, ISBN 978-99955-36-07-7

[2] T. Jiahua, Z. Du, S. Xibin and Q., “Improving Temporal Event Scheduling through STEP Perpetual Learning“, *Sustainability* 2022, 14(23), 16178; https://doi.org/10.3390/su142316178

[3] M. I. Jordan, T. M. Mitchell, „Machine Learning: Trends, Perspectives, and Prospects“, *Science* 2015, 349, 255–260.

[4] M. Banjanin, „Simetrične kombinacije multi-modalnih interakcija u inteligentnoj okolini“, Fakultet tehničkih nauka: Centar za razvoj i primenu nauke, tehnologije i informatike: Prirodno matematički fakultet; Beograd: Institut za pedagoška istraživanja, (Čačak: Fakultet tehničkih nauka), UDK COBISS.SR-ID 250182407, ISBN ISBN 978-86-7447-090 No. 5, Str. 124-135, Beograd, 2010

[5] B. Žanić, „U perspektivi dinamične semantike: valjanost praktičnog zaključka“, University of Split, January 2005.

[6] B. Žarnić, “Imperative negation and dynamic semantics”, J. Peregrin (ured.) Meaning: the Dynamic Turn. str. 201–211. Current Research in the Semantics/Pragmatics Inter-face: Volume 12. Elsevier Science, Oxford, 2003.

[7] D. L. Silver, Q. Yang, L. Li, „Lifelong Machine Learning Systems: Beyond Learning Algorithms“, *In Proceedings of the 2013 AAAI Spring Symposium* Series, Palo Alto, CA, USA, 25–27 March 2013.

[8] B. A. Buckler, „Learning Process Model to Achieve Continuous Improvement and Innovation“, The Learning Organization: Bingley, UK, 1996.

[9] M. Banjanin, „Naučnoistraživačka metodologija“, DisPublic, Beograd 2008, (stranica 480), 2. prerađeno izdanje i dopunjeno izdanje, ISBN 86-86035-01-9, COBISIS:SR-ID 128215308

[10] T. Fourez, N. Verstaevel, F. Migeon, F. Schettini and F. Amblard, „An ensemble Multi-Agent System for non-linear classification“, *arXiv preprint arXiv:2209.06824*, 2022

[11] https://www.projectpro.io/article/bagging-vs-boosting-in-machine-learning/579 (01.05.2023.)

[12] A. Đukić, R. Bjelošević, M. Stojčić, M. K. Banjanin, “Network Model of Multiagent Communication of Traffic Inspection for Supervision and Control of Passenger Transportation in Road and City Traffic”, Croatian Society for Information, Communication and Electronic Technology – MIPRO, MIPRO 2023 - 46th Convention, Croatia, Opatija, 22 - 26 May 2023 (paper accepted).

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

[14] IBM, “IBM SPSS Modeler 18.0 User's Guide” https://public.dhe.ibm.com/software/analytics/spss/documentation/modeler/18.0/en/ModelerUsersGuide.pdf (01.05.2023)