MAGYAR AGRÁR- ÉS ÉLETTUDOMÁNYI EGYETEM

Application of artificial intelligence methods in the security audit of IT systems

DOI: 10.54598/001400

Gergő Barta

Gödöllő

2021

The doctoral school

name: Doctoral School of Economics and Regional Science

scientific scope: economics and regional science

head: Prof. Dr. H. c. Popp, József

Corresponding member of the Hungarian Academy of Sciences

Hungarian University of Agriculture and Life Sciences

Director of Institute of Economic Sciences

Head of the Doctoral School of Economic and Regional Sciences

Témavezető: Dr. László Pitlik

Associate Professor, Head of Department

Kodolányi János University

Institute of Sustainable Economy

Department of Informatics

....

Approval of the head of the doctoral school

Approval of the Supervisor

Table of contents

1. BACKGROUND AND OBJECTIVES

There is a constant and growing demand for business process automation (STEGMAN et al. 2019), which today can no longer be considered exclusively as the automation of routine work or rule-based activities, but also of more complex tasks that require creativity and even subjective value judgment. Today's technologies, such as Big Data, Artificial Intelligence, Cloud-Based Solutions, and the growth of hardware resources (e.g., HPC - High Performance Computing) and the gradual improvement of their performance, enable high-quality automated decision-making. Its professional and research-intensive area has more and more encouraging results year by year (e.g. BODA 2019).

However, the development of integrated, IT-enabled solutions to support automation and the modernization of obsolete systems impose challenges for technicians, managers, and users, as they carry risks (e.g., breach of confidentiality of business data assets stored in cyberspace) and contributes to the birth of functional units dealing with information management, security and control - creating a risk management process for IT resources (BARTA - GÖRCSI 2021).

The final outcome of risk analysis procedures is a strategic decision to reduce the IT risks in the organization (VASVÁRI 2008). A decision based on the results of the IT risk analysis may be to take measures to mitigate the risks that arise, i.e. to mitigate the identified risks through control processes, such as the implementation of preventive precautions, corrective or detective activities.

The existence of implemented IT controls does not guarantee fault-free operation, i.e. the management must regularly make sure that the mitigating measures work effectively, it is not possible to circumvent them intentionally (POMPON 2016).

The functional area for exploring the effectiveness of internal IT controls is the IT audit team, which consists of professionals with business and IT expertise whose primary goal is to review and continuously test the internal IT control environment to ensure that IT processes are in alignment with organizational business objectives while maintaining the highest possible level of information security. The emergence and introduction of auditing in an organization may help to identify control gaps, but there is still the human factor, i.e. the possibility of accidental and / or intentional error, and human

resource constraints (BARTA 2018b). In most cases, it is almost impossible to perform the control testing manually because of its dynamic characteristics.

For this reason, auditors try to filter out control deficiencies with a predetermined number of statistical sample items, i.e., the audit is only partial (BARTA 2018b). In addition, there are many examples in Hungary and from abroad where the quality of audit work has failed due to conflicts of independence, conflicts of organizational interests or unethical business conduct:

- Think of the 2001 Enron scandal, which resulted in the dissolution of Arthur Andersen, the world's 5th largest audit firm, as held accountable by the U.S. Supreme Court for its professionally and morally objectionable practices (GREENHOUSE 2005).
- A negative example can also be mentioned in Hungary, which can be related to the negligence of IT controls:
    - On 24 February 2015, the Hungarian National Bank (MNB) suspended the operation of Buda-Cash Bróker Zrt. (Magyar Nemzeti Bank 2015),
    - which led to the fall of Hungária Securities Ltd. and
    - Quaestor Securities Trading and Investment Ltd., which together issued approximately HUF 250 billion in fictitious bonds.

The negative examples listed above and the risks associated with the operation of audit, can make it clear that the topic is relevant and actual, and will become even more prominent and significant in the near future (c.f. Industry 4.0, MI Coalition), as the business benefits of IT industrialization and business automation will enforce competitors to implement technology solutions (VINOGRADOV 2020).

In order to increase the efficiency of audit work, an automated solution is needed that is able to objectively search for signs of deficiencies, process and interpret them, not only performing rule-based evaluation, but also explore complex, underlying interrelations that can be treated effectively. Based on my, the practical enforcement of these mentioned expectations can be included in the concept of artificial intelligence. Artificial intelligence is a raison d'être to solve the problem by following Knuth's (1995) principles: "*Science is what we understand well enough to explain to a computer; art is everything else*." This shall be the basic expectation of the 21st century (PITLIK et al. 2017).

Artificial intelligent enabled software solutions e.g. through their learning mechanisms, may be able to monitor suspicious activities, thereby effectively creating real value in IT auditing. The problem raised is, therefore, the decision situation-specific development of the concept of suspicion generation to detect control deficiencies using artificial intelligences, i.e. the replacement and support of living human capabilities.

One of the biggest challenges of artificial intelligence software solutions is the limited availability and diversity of samples used to teach and test systems, so it is necessary to research methods that can exploit the maximum potential (information added value) in the available data set. One of the central results of the dissertation is to find the optimum in the available data, on the one hand, through more efficient processing of the training data set, and through the derivation of model preference without classical testing procedures.

The objectives and hypotheses of the research can be structured according to the following points:

C1: Following the Knuth principle, in order to increase the efficiency of information security audits, a decision support system (robot auditor) should be created with artificial intelligence, which can automatically explore the interrelations between control gaps and control areas and make suggestions for detection of potential human error to reduce detection risk.

H1: Creating a structured database from information security audit reports and processing it as input into the decision support system, the existence of control deficiencies to be detected during audits is more likely to be detected than random guessing.

- **H1.1**: Evaluating the specifics of suspicion generation as a business-interpreted problem to be solved, the detection of control deficiencies can be solved by both supervised and unsupervised machine learning procedures.

- **H1.2**: The performance of suspicion generation can be enhanced in a hybrid approach, i.e. the relevant performance metrics used in the research of the combined use of supervised and unsupervised methods show more ideal values than in stand-alone application.

- **H1.3**: The hybrid model is able to improve the generalization power of simple models.

C2: The software robot auditor with artificial intelligence to be developed, must be able to use the available data set as optimally as possible, thereby maximizing its performance, i.e. the goal is to exploit the genetic potential of the robot auditor through controlled reduction of the learning data set.

H2: The genetic potential of a decision support system can be reached by a search procedure with similarity analysis through controlled processing of the data set used for training, so that the search procedure leading to genetic potential can provide a more ideal result without random mutation and population crossings.

C3: In order to increase the information value of the data set used for machine learning, it is necessary to compare the performance of each robot auditor alternative in an anti-discriminatory way, to select the best alternative, which does not require separation of validation and test set.

H3: Artificial intelligence decision support systems can be ranked on a performance basis without classical testing procedures for machine learning applications, deriving the value direction of the descriptive properties of predictions as generated sources of suspicion and the mathematical apparatus processing this data automatically to objectively determine preferred models.

2. MATERIAL AND METHODS

The research presented in the dissertation required thorough planning work, therefore its process was structured according to pre-defined professional and scientific methodological guidelines, focusing on the respective objectives. At the strategic level of the design of the research, the most important thing was to be observed as a minimum requirement for all artificial intelligence-based projects: the concept of "good" should be defined in advance and as accurately as possible algorithmically based on Knuth (1995).

2.1. Data collection

To perform the research work, the corresponding task was to design databases, where the critical element was the realistic, reliable and objective compilation of data that are already able to achieve the modeling objectives, i.e. give a clear picture of which information security control areas pose the greatest risk.

Since information security exposures, lack of controls, and their improper operation are confidential information, and their disclosure is a vulnerability for most organizations, companies are in most cases unwilling to declare their adequacy, or obscuring the whole truth, they are to state controversial information to report even within their organization because of the fear of the consequences of disciplinary actions.

The source of the data used for the research contains real data assets collected during an on-site inspection (fieldwork) from two auditors, advisory and consulting organizations - each with obtained consent, employing more than 500 employees.

The audit organizations provided me with a total of 127 real audit reports in textual form, in an anonymised manner about audits completed between 2016-2020. The audit reports represent 127 different independent audits. The shortest audit report was 5 pages, while the longest was 106 pages. Data were collected between 1 April 2020 and 30 June 2020 and a total of 3699 pages of audit reports were processed (average rounded to 29 pages per audit report).

It should be noted that data collection based on questionnaires and / or interviews was deliberately excluded during the preparation of the research work, as managers do not intend to share control gaps with external parties, and only independently collected data is objectively reproducible describing reality in a verifiable way.

The raw attributes of the database were defined according to Annex A of ISO / IEC 27001: 2013.

2.2. Applied algorithms and statistical procedures

The following subsection provides the modeling practices used to prove the hypotheses, the input of which was the data assets discussed in the previous subsection, as well as the related calculation results.

Decision tree

Decision tree-based procedures formulate a series of questions based on a selected attribute of the data set, where each subsequent question refers to a threshold (condition) for a particular attribute. The decision tree was used in the development of the decision support system (robot auditor).

Neural network

The theoretical approach to neural networks is the embodiment of the biological neurons of the human brain into a machine form, for which it also inherited its name (KÁSA 2018). Similar to the structure of natural neurons, artificial neurons are interconnected, where they transport the processed information from neuron to neuron. The neurons are located in different layers: the input layer receives the data needed for processing, the output layer provides the prediction, and the hidden layers transform the information from the input neuron to the output neuron.

The transferred information is "transformed" on the way from one neuron to another through the mediation of weight matrices and activation functions. The weight matrices contain the "knowledge" of the neural network, and the neural network modifies it according to the given learning procedure and the pattern found in the data. The neural network calculates the difference between the factual value of the target variable and the estimate, according to the defined loss function, which it decrypts (e.g. backpropagation) to modify the weight matrices of the network (gradient procedure) (BARTA 2018a, BARTA - PITLIK 2018). The neural network was used in the development of the decision support system (robot auditor).

Adaptive Boosting (ABM)

Adaptive Boosting (AdaBoost or ABM - Adaptive Boosting Machine) is an ensemble method of re-weighting incorrectly classified records using several weak learners to form a strong classifier. ABM is a sequential method, so each weak learner builds on each other (does not operate independently), with an emphasis on input data. Each set of data used for learning depends to some extent on the previous weak learner, gradually correcting the mistakes made by the previous classifiers. The ABM combines individual predictions of weak learners as a final classification model, where it takes the mode of partial results (majority voting). It is by default a method used to ensemble decision trees, but due to its characteristics it can be used for sequential connection of any classifier. ABM was used in the development of the decision support system (robot auditor).

Gradient Boosting (GBM)

Gradient Boosting (Gradient Boosting Machine or GBM) is a sequential method similar to the ABM algorithm, however, it is a fine-tuned variant of ABM, but it seeks to create a strong classification algorithm by optimizing the differentiable loss function (gradient) (MASON et al. 1999, FRIEDMAN 2001). GBM was used in the development of the decision support system (robot auditor).

Collaborative filtering

The method is able to identify similar audit reports with distance / relationship metrics, so it is a special clustering procedure where the most similar objects return a list of suspected suspicious controls. The list of the two e.g. gives the differences between the objects considered to be most similar as output, where the degree of similarities between the objects can be used as a weight number. It is the decision maker’s preference to determine the threshold for acceptable similarity, or the number of minimum and maximum audit reports that are closest to the selected object in the multidimensional space. Collaborative filtering was used in the development of the decision support system (robot auditor).

Similarity analysis

One method of similarity analysis is the mathematical embodiment of the "everyone is alike" principle, in which the goal of the modeling is to find the most ideal object by defining a fictitious target variable that value is constant and which is ranked according to the given direction preferences. Using a constant target variable, object identities appear as constraints in the modeling process, which is the central element (norm value) of anti-discriminative calculations, i.e., the similarity scale.

Similarity analysis is an optimization procedure that approximates a stepwise function for each attribute, which determines the extent of the contribution to the target variable (BÁNKUTI 2010). The minimum differences among the steps (ideality weights) must always be greater than zero, so different rankings must have different step values. By similarity analysis, thus, objects significantly different from the norm, e.g. appear as outliers, can be ranked according to the value of the target variable. The anti-discriminatory method of similarity analysis can be considered as a special neural network, methodologically unsupervised machine learning algorithm.

2.3. Evaluation metrics

To determine the objective results of the research and to measure the performance of the models, I use the following metrics:

- **Number of true positives**: The number of true hits for the model when it correctly hit the suspicious control that occurred.
- **Number of true negatives**: The number of true hits for the model when it correctly found that a particular control is not a suspicious case.
- **Number of false positives**: The number of false positives in the model when it incorrectly concluded that the particular control was a suspicious case and not in fact.
- **Number of false negatives**: The number of false hits in the model when it incorrectly concluded that the particular control was not a suspicious case, and in fact, it was.
- **Accuracy**: The quotient of the sum of True Positives and True Negatives and all cases.
- **Precision**: Precision is the ratio of the number of True Positives to the sum of True Positives and False Positives. In the case of imbalanced

data sets, the preferred indicator for the relevance of the model accuracy estimate is "How valid is the result?".

- **Recall**: Recall is the ratio of the number of True Positives to the sum of True Positives and False Negatives. In the case of imbalanced data sets, the preferred indicator for the relevance of the model accuracy estimate is "How complete is the result?".
- **F1-Score**: F1-Score is the harmonic mean of the performance derived from the combination of Precision and Recall.
- **Variance**: The model variance index expresses the difference between the accuracies measured on the learning and test data set.
- **ROC curve and AUROC**: The application of the method makes it possible to explore and interpret the trade-off between the ratios of true positive and false negative hits in the classification task.
- **PR curve and AUPRC**: It illustrates the trade-off between the Precision and Coverage metrics, similar to the ROC curves.

3. RESEARCH RESULTS

The chapter presents the proof of the hypotheses set up based on the analysis of the data collected in the fieldwork using the mathematical apparatus described in the previous chapter.

3.1. Suspicion generation to detect information security control deficiencies

3.1.1. The development and evaluation of supervised learning models

In the modeling phase of the supervised machine learning algorithms, I processed the transformed data set with three different methods:

• ABM (Adaptive Boosting) using decision tree;

• GBM (Gradient Boosting) using decision tree;

• NN (Neural Network).

After the development and "deployment" of the algorithms, I recorded the results described in Table 1, measured on the test set, which were provided by the cross-validation evaluation.

ABM (80), then NN (74), and subsequently GBM, which was only able to perform true positive hits in 46 cases, thus remaining below the performance of the other two methods used, produced most true positive hits. However, examining the number of true negatives, GBM (954) outperformed the ABM (919) and NN (916) algorithms, which presumably supports that GBM rated samples (controls) as appropriate by default, with low false positives. The F1-Score expresses the combined harmony of Recall and Precision on a scale from 0 to 1 (multiplied by 100 to obtain percentage performance, similar to Accuracy, Precision, Recall, and AUROC), which for ABM (0.57) is the most favorable. The AUROC indices of ABM and GBM (0.85) are the most ideal, and the GBM algorithm has the least variance (0.04).

Table 1: Performance metrics of supervised algorithms

KPIs	ABM	GBM	NN
True Positives	80	46	74
True Negatives	921	954	916
False Positives	42	9	47
False Negatives	79	113	85
Accuracy	0.89	0.89	0.88
Precision	0.66	0.83	0.61
Recall	0.50	0.29	0.47
F1-Score	0.57	0.43	0.53
Variance	0.08	0.04	0.11
AUROC	0.85	0.85	0.82
AUPRC	0.58	0.61	0.56

Source: Own editing

In general, it can be observed that all three methods are overfit (variance> 0), i.e. their generalization power is not optimal, the accuracy measured on the learning set exceeds the result of the test set, so the algorithms have also incorporated the noise into the approximation that was expected.

The value of the variance is the smallest in the case of GBM (0.04), while the highest in the NN model (0.11), where learning is almost at maximum around 2,000 samples (the algorithm converged), so the model correctly evaluated all samples in the learning set, which is high signs of overfitting and low generalization power is presumed, so the sustainability of the application in a production environment can be questioned. The degree of overfit can be reduced by obtaining additional learning data as well as by regularizing the models.

Examining the AUROC results, it can be observed that all three methods were able to mathematically map the pattern between control deficiencies as they performed above the desired random level (AUROC value exceeded 0.5). Furthermore, since the value of F1-Score can be interpreted and its value is greater than zero (the algorithms did not classify all records into a given class), it can be stated that the performance of the systems is better than random guessing i.e. interrelation exists.

3.1.2. The development and evaluation of unsupervised learning models

Unsupervised methods can also be used to generate suspicion, however, measuring performance is cumbersome because there is no confirmed target variable due to the specifics of the procedures. The decision-making system used for suspicion generation can also be interpreted as a recommendation system, where the system must artificially "recommend" suspicious controls based on historical data.

In the experiment, I used collaborative filtering, where I used three different distance / relationship metrics for the algorithms, so three different models were generated: EUC, PEA and COS.

Twenty-five random samples were selected for testing, representing 19.69% of the total population (127). After running the algorithms, I recorded the results described in Table 2.

Table 2: Performance metrics of unsupervised models

KPIs	EUC	PEA	COS
True Positives	16	18	20
True Negatives	2,739	2,721	2,710
False Positives	111	129	140
False Negatives	9	7	5
Accuracy	0.96	0.95	0.95
Precision	0.13	0.12	0.13
Recall	0.64	0.72	0.80
F1-Score	0.21	0.21	0.22

Source: Own editing

For the selected sample, depending on the methodology used, the algorithms were able to form a true positive opinion on a total of potentially 25 suspicious controls about the putative control deficiencies. Of this, 80% of COS (20), 72% of PEA (18) and 64% of EUC (16) were correctly judged with a high false positive hit rate.

The Recall, which expresses the ratio of true and false positives, shows a favorable value for all three models due to the use of an artificially enforced latent target variable. Since the return lists contain all the discrepancies that occur between the selected objects, the low level of Precision is not surprising.

The nearly identical value of F1-Score does not allow to conclude a significant difference between the procedures, with minimal difference, COS can be considered the best method from the aspect of F1-Score.

3.1.3. Suspicion generation with hybrid approach

The basic utilization of hybrid modeling is to increase the performance metrics using the unsupervised algorithms described in the previous subsection, thus incorporating "recommendations" into supervised methods to provide additional useful input, presumably creating a better model.

In the hybrid modeling, I incorporated the output values of the COS model into the supervised methods (ABM, GBM, NN) ceteris paribus, preserving all the configuration settings of the supervised methods, thus making changes (extensions) to the learning set only. I ran the COS model on all audit reports, providing input for both the training and test set. Out of the possible 115 controls, the model returned with suspicion in 98 cases, so a total of 98 binary variables were added to the learning set, which contains the recommendations of the COS model. Because the configuration settings for the procedures are identical, the improvement / deterioration effects of the hybrid solution can be objectively verified.

Running the hybrid model in production environment, the performance metrics of the supervised methods were modified as follows (Tables 3., 4. and 5). I called methods without a hybrid approach "simple" models.

Table 3: Performance metrics of simple and hybrid ABM models

KPIs	Simple ABM	Hybrid ABM	Change (%)	Objective	Impact of the change
True Positives	80	89	+11.25%	increase	improvement
True Negatives	921	944	+2.50%	increase	improvement
False Positives	42	19	-54.76%	decrease	improvement
False Negatives	79	70	-11.39%	decrease	improvement
Accuracy	0.89	0.92	+3.37%	increase	improvement
Precision	0.66	0.82	+24.24%	increase	improvement
Recall	0.50	0.56	+12.00%	increase	improvement
F1-Score	0.57	0.67	+17.54%	increase	improvement
Variance	0.08	0.06	-0.25%	decrease	improvement
AUROC	0.85	0.90	+5.88%	increase	improvement
AUPRC	0.58	0.70	+20.69%	increase	improvement

Source: Own editing

Table 4: Performance metrics of simple and hybrid GBM models

KPIs	Simple GBM	Hybrid GBM	Change (%)	Objective	Impact of the change
True Positives	46	62	+34.78%	increase	improvement
True Negatives	954	958	+0.42%	increase	improvement
False Positives	9	5	-44.44%	decrease	improvement
False Negatives	113	97	-14.16%	decrease	improvement
Accuracy	0.89	0.91	+2.25%	increase	improvement
Precision	0.83	0.93	+12.05%	increase	improvement
Recall	0.29	0.39	+34.48%	increase	improvement
F1-Score	0.43	0.55	+27.91%	increase	improvement
Variance	0.04	0.03	-0.25%	decrease	improvement
AUROC	0.85	0.85	0%	increase	neutral
AUPRC	0.61	0.71	+16.39%	increase	improvement

Source: Own editing

Table 5: Performance metrics of simple and hybrid NN models

KPIs	Simple NN	Hybrid NN	Change (%)	Objective	Impact of the change
True Positives	74	76	+2.70%	increase	improvement
True Negatives	916	926	+1.09%	increase	improvement
False Positives	47	37	-21.28%	decrease	improvement
False Negatives	85	83	-2.35%	decrease	improvement
Accuracy	0.88	0.89	+1.14%	increase	improvement
Precision	0.61	0.68	+11.48%	increase	improvement
Recall	0.47	0.47	0%	increase	neutral
F1-Score	0.53	0.56	+5.66%	increase	improvement
Variance	0.11	0.11	0%	decrease	neutral
AUROC	0.82	0.85	+3.66%	increase	improvement
AUPRC	0.56	0.62	+10.71%	increase	improvement

Source: Own editing

The hybrid modeling has clearly improved the results, with almost without exception an improvement in all metrics can be observed.

3.2. Searching for genetic potential using reduced use of the data set

The genetic potential determined by the optimal use of the training set of the learning algorithm should be approximated by a search algorithm in which the search is successful if an objectively recognizable, more ideal result is obtained in one or more steps, so that the genetic potential can be identified as an outlier.

The search procedure I developed, improved the performance of the hybrid ABM, GBM and NN algorithms as can be observed in the following tables (Tables 6, 7 and 8), where the maximum was found in the first iteration in the case of ABM, while in the case of the NN method, it was occurred in the second iteration, with minimal improvement. It should be emphasized that heuristics do not necessarily find the best possible solution, as the quality of the search is fundamentally affected by the initialization of the base population.

Table 6: ABM performance metrics by using search algorithm

KPIs	Simple ABM	Hybrid ABM	Change (%)	Objective	Impact of the change
True Positives	89	102	+14.60%	increase	improvement
True Negatives	944	940	-0.42%	increase	deterioration
False Positives	19	23	+21.05%	decrease	deterioration
False Negatives	70	57	-18.57%	decrease	improvement
Accuracy	0.92	0.93	+1.09%	increase	improvement
Precision	0.82	0.82	0%	increase	neutral
Recall	0.56	0.64	+14.29%	increase	improvement
F1-Score	0.67	0.72	+7.46%	increase	improvement
Variance	0.06	0.06	0%	decrease	neutral
AUROC	0.90	0.86	-4.44%	increase	deterioration
AUPRC	0.70	0.71	+1.43%	increase	improvement

Source: Own editing

Table 7: GBM performance metrics by using search algorithm

KPIs	Simple GBM	Hybrid GBM	Change (%)	Objective	Impact of the change
True Positives	62	78	+25.81%	increase	improvement
True Negatives	958	955	-0.31%	increase	deterioration
False Positives	5	8	+60.00%	decrease	deterioration
False Negatives	97	82	-15.46%	decrease	improvement
Accuracy	0.91	0.92	+1.10%	increase	improvement
Precision	0.93	0.91	-2.15%	increase	deterioration
Recall	0.39	0.48	+23.08%	increase	improvement
F1-Score	0.55	0.64	+16.36%	increase	improvement
Variance	0.03	0.03	0%	decrease	neutral
AUROC	0.85	0.85	0%	increase	neutral
AUPRC	0.71	0.68	-4.22%	increase	deterioration

Source: Own editing

Table 8: NN performance metrics by using search algorithm

KPIs	Simple NN	Hybrid NN	Change (%)	Objective	Impact of the change
True Positives	76	80	+5.26%	increase	improvement
True Negatives	926	926	0%	increase	neutral
False Positives	37	37	0%	decrease	neutral
False Negatives	83	79	-4.82%	decrease	improvement
Accuracy	0.89	0.90	+1.12%	increase	improvement
Precision	0.68	0.68	0%	increase	neutral
Recall	0.47	0.50	+6.38%	increase	improvement
F1-Score	0.56	0.58	+3.57%	increase	improvement
Variance	0.11	0.11	0%	decrease	neutral
AUROC	0.85	0.81	-4.71%	increase	deterioration
AUPRC	0.62	0.62	0%	increase	neutral

Source: Own editing

The tables illustrate that some of the indicators in the algorithms deteriorated in order to improve the F1-Score, as the search procedure was used to scan the more ideal value of the F1-Score by ignoring the other metrics. Without exception, this was achieved by improving the rate of true positive (and thus false negative) hits. The number of true negative and false positive hits for ABM and GBM-based systems decreased, while it did not change for NN. AUROC indicators deteriorated with the exception of GBM, while variance

was not affected by the procedure, i.e., the accuracy of learning improved at the same rate as that of testing. Overall, it can be stated that the search procedure was successful, for all three algorithms the F1-Score was improved by using the training set closer to the optimum.

3.3. Derivation of model preference without classical testing procedures

The intention of model preference measurement is to create a data set from the descriptive properties of suspicious objects generated by the system, in which it is possible to determine along several dimensions which system responses are preferred (e.g. by the current decision maker or the decision support system itself) more than another. Thus, what are the characteristics of a system response that can provide higher certainty (cf. consistency (PITLIK et al. 2017)) to the user of the decision support system. An application that generates a large amount of suspicion about different controls, different control areas, is likely to destabilize the decision maker and may lead to a higher false positive result.

An application that is able to point in one direction in a uniform and homogeneous way (e.g., a control is marked as a suspicious case or suggests a set of controls with a similar structure / function for review) is likely to be easier to accept because each component in the system has a similar conclusion.

The descriptive properties to be defined, therefore, the widest possible aspects of successful and unsuccessful modeling, and should be defined in the form of descriptive measures of model performance. Since the number of these aspects is infinite, the classical one-factor optimization is forced to be replaced by a multidimensional value concept to fine-tune the Hartman principle (PITLIK et al. 2020).

By developing my own model preference derivation algorithm, I examined the behavior of 12 models in the case of unsupervised machine learning technics. Performance metrics are illustrated in Table 9.

Table 9: Summary table of usupervised model evaluation

Model ID	Naïve Average	Ranking	Similarity ideals	Ranking by similarity analysis
R_NS_NRS_C	5.75	6	997	8
R_NS_NRS_E	**3.00**	**1**	**1012**	**1**
R_NS_NRS_P	5.75	6	1002	4
R_NS_RS_C	5.75	6	997	8
R_NS_RS_E	4.75	3	991	11
R_NS_RS_P	5.75	6	1002	4
R_S_NRS_C	6.00	10	996	10
R_S_NRS_E	3.75	2	1003	3
R_S_NRS_P	5.00	5	1008	2
R_S_RS_C	6.25	11	1002	4
R_S_RS_E	4.75	3	991	11
R_S_RS_P	6.25	11	1002	4

Source: Own editing

Based on the table, it can be concluded that R_NS_NRS_E is the most ideal model for solving the raised suspicion generation problem among all the developed models, which is supported by both rankings (naive averaging and similarity analysis).

After evaluating the test set in the simulated environment and comparing the results with those described, the following table illustrates the summary results for the hit rates of the models (Table 10).

Table 10: Evaluation matrix of unsupervised models

Model ID	Accuracy	Precision	Recall	F1-Score
R_NS_NRS_C	0.78	0.12	0.42	0.14
R_NS_NRS_E	0.84	0.24	0.67	0.32
R_NS_NRS_P	0.83	0.13	0.42	0.16
R_NS_RS_C	0.78	0.12	0.42	0.14
R_NS_RS_E	0.61	0.22	0.92	0.30
R_NS_RS_P	0.83	0.13	0.42	0.16
R_S_NRS_C	0.75	0.11	0.42	0.13

R_S_NRS_E	0.81	0.22	0.67	0.30
R_S_NRS_P	0.80	0.12	0.42	0.14
R_S_RS_C	0.74	0.11	0.42	0.13
R_S_RS_E	0.56	0.18	0.92	0.25
R_S_RS_P	0.79	0.12	0.42	0.15

Source: Own editing

Performing the evaluation of the applied performance metrics (Accuracy, Precision, Recall, F1-Score) by the anti-discriminatory mathematical apparatus provided by the similarity analysis, where the value of the norm is 1,000 (Table 11), it can be stated that the indicators of the R_NS_NRS_E model are the most ideal overall (1,016.60), so the result of the feedback is that the model was rightly declared the winner.

Table 11: Model ideality prediction by similarity analysis

ID	Model ID	Accuracy	Precision	Recall	F1-Score	Y_0
1	R_NS_NRS_C	0.78	0.12	0.42	0.14	**995.60**
2	R_NS_NRS_E	0.84	0.24	0.67	0.32	**1016.60**
3	R_NS_NRS_P	0.83	0.13	0.42	0.16	**1005.60**
4	R_NS_RS_C	0.78	0.12	0.42	0.14	**995.60**
5	R_NS_RS_E	0.61	0.22	0.92	0.30	**1003.60**
6	R_NS_RS_P	0.83	0.13	0.42	0.16	**1005.60**
7	R_S_NRS_C	0.75	0.11	0.42	0.13	**986.60**
8	R_S_NRS_E	0.81	0.22	0.67	0.30	**1011.60**
9	R_S_NRS_P	0.80	0.12	0.42	0.14	**997.60**
10	R_S_RS_C	0.74	0.11	0.42	0.13	**985.60**
11	R_S_RS_E	0.56	0.18	0.92	0.25	**998.60**
12	R_S_RS_P	0.79	0.12	0.42	0.15	**997.60**

Source: Own editing

Table 9. showed the ranking of the models derived using the object-descriptive properties. Calculating the correlation with the similarity analysis with idealities (Y_0) described in Tables 35 and 11, the value of the correlation coefficient is 0.42, which assumes a medium positive relationship. It should be emphasized that the model preference for both procedures was R_NS_NRS_E, and the derivation efficiency was independently confirmed.

4. NEW AND NOVEL SCIENTIFIC RESULTS

As a result of my research, I recorded the new and / or novel scientific results listed below in accordance with the hypotheses:

H1*: H1: Creating a structured database from information security audit reports and processing it as input into the decision support system, the existence of control deficiencies to be detected during audits is more likely to be detected than random guessing.* **Verified**

- Audit findings, i.e. control deficiencies, which are communicated in the form of a textual report to managers and investors of organizations, can be categorized in a logical context-dependent form, and upload to the machine in a structured form after conscious systematization of the findings;

- There is an interrelation between the joint constellations of the findings documented by the audits, which can be demonstrated objectively, using mathematical apparatus, so a decision support system can be built, which is able to increase the efficiency of audits by ensuring their objective quality assurance:
    - Simple and hybrid ABM F1-Score values: 0.57 and 0.67, respectively;
    - Simple and hybrid GBM F1-Score values: 0.43 and 0.55, respectively;
    - Simple and hybrid NN F1-Score values: 0.53 and 0.56, respectively;

- Based on the co-existence of control deficiencies, control deficiencies can be estimated for a given control on data not processed in advance by the decision support system, a more ideal result can be achieved than random guessing (AUROC> 0.5):
    - Simple and hybrid ABM AUROC values: 0.85 and 0.90, respectively, AUPRC values: 0.58 és 0.70, respectively;
    - Simple and hybrid GBM AUROC values: 0.85 and 0.85, respectively, AUPRC values: 0.61 és 0.71, respectively;
    - Simple and hybrid NN AUROC values: 0.82 and 0.85, respectively, AUPRC values: 0.56 és 0.62, respectively;

***H1.1**: H1.1: Evaluating the specifics of suspicion generation as a business-interpreted problem to be solved, the detection of control deficiencies can be solved by both supervised and unsupervised machine learning procedures.* **Verified**

- Identifying the detection of control deficiencies as a classification problem, the generation of suspicion can be implemented in the framework of supervised machine learner modeling, where the dedicated target variable indicates the state of compliance of the given control:
    - Simple ABM F1-Score value: 0.57;
    - Simple GBM F1-Score value: 0.43;
    - Simple NN F1-Score value: 0.53;

- Identifying the detection of control deficiencies as a recommendation system, suspicion generation can be implemented in the framework of unsupervised machine learner modeling, which, in the absence of a target variable, returns a list of identified suspicion moments to be discovered based on audit report similarity:
    - EUC F1-Score value: 0.21;
    - PEA F1-Score value: 0.21;
    - COS F1-Score value: 0.22;

***H1.2**: H1.2: The performance of suspicion generation can be enhanced in a hybrid approach, i.e. the relevant performance metrics used in the research of the combined use of supervised and unsupervised methods show more ideal values than in stand-alone application.* **Verified**

- The performance of detecting control deficiencies can be improved by the combined use of supervised and unsupervised methods, where the unsupervised recommendation system suggests the moments of suspicion that the supervised methods incorporate into decision-making, thus reducing the degree of uncertainty as additional information:
    - Simple and hybrid ABM F1-Score values: 0.57 and 0.67, respectively;
    - Simple and hybrid GBM F1-Score values: 0.43 and 0.55, respectively;
    - Simple and hybrid NN F1-Score values: 0.53 and 0.56, respectively;

***H1.3**: H1.3: The hybrid model is able to improve the generalization power of simple models.* **Partly Verified**

- The variances of supervised machine learning systems are reduced by the hybrid approach, so their generalization power is more ideal than in stand-alone application:
    - Simple and hybrid ABM Variance values: 0.08 and 0.06, respectively;
    - Simple and hybrid GBM Variance values: 0.04 and 0.03, respectively;
    - Simple and hybrid NN Variance values, respectively: 0.11 and 0.11 (variance did not decrease for neural network);

***H2**: The genetic potential of a decision support system can be reached by a search procedure with similarity analysis through controlled processing of the data set used for training, so that the search procedure leading to genetic potential can provide a more ideal result without random mutation and population crossings.* **Verified**

- The step functions generated by the similarity analysis are suitable for use in search procedures;

- The genetic potential of the decision support system can be sought by deriving the value direction of the rational descriptive attributes belonging to the training set;

- The search procedure can determine, based on the values of the available population, whether a given model is likely to have reached its genetic potential through an attempt to optimize the training set, and what attributes need to be modified to move toward the ideal target variable:
    - The algorithm gives the genetic potential of the model in the light of a given target variable as the sum of the most prominent weights (for GBM in iteration 11, the genetic potential of F1-Point: 0.71);
    - The algorithm determined, due to the difference in weights, which attributes needed to be modified to move in the direction of the ideal target variable (e.g., for GBM, it was achieved in iteration 1 by increasing the f_{10} attribute according to a specific direction preference, where it was 1.9).

- The search procedure is able to indicate a more ideal direction even without accidental mutation:
    - The algorithm improved the GBM F1-Score value from 0.60833 to 0.61925 already in iteration 1 without random mutation, which meant an improvement of 1.80%;
    - The algorithm improved the GBM F1-Score value from the initial 0.60833 to 0.63673 in iteration 11 without random mutation, which represented a 4.57% improvement.

- The search procedure can indicate a more ideal direction without crossing the individuals:
    - The algorithm improved the value of GBM F1-Score from 0.60833 to 0.61925 already in iteration 1 without crossing the individuals, which meant an improvement of 1.80%.
    - The algorithm improved the GBM F1-Score value from the initial 0.60833 to 0.63673 in iteration 11 without crossing the individuals, which represented an improvement of 4.57%.

- The search procedure does not require intermediate new populations beyond the opening population;

- The directions can be parameterized automatically in each iteration.

H3*:Artificial intelligence decision support systems can be ranked on a performance basis without classical testing procedures for machine learning applications, deriving the value direction of the descriptive properties of predictions as generated sources of suspicion and the mathematical apparatus processing this data automatically to objectively determine preferred models.*.
Verified

- The developed method for searching for model preference is suitable for finding the ideal model from a predefined set of models, which can be objectively proved by independent testing:
    - The R_NS_NRS_E model was the most ideal, with a similarity analysis ideality of 1.012 (highest among all models), supported by an independently measured model performance estimate (similarity analysis ideality: 1.017, highest among all models).

- There is a correlation between the self-reinforcing (consistency) mechanisms of the models (behavior of system responses) and model performance metrics, so through the discovery of inconsistencies, performance can be estimated:
    - Calculating the correlation with the similarities of the similarity analysis with ideals (Y_0), the value of the correlation coefficient is 0.42, which assumes a medium positive relationship.

- Determining the direction preference of the attributes belonging to the system responses communicated by the decision support system can be objectified, so it is possible to minimize human errors in the decision-making process;

5. CONCLUSIONS AND RECOMMENDATIONS

Based on the research conducted, the following points summarize the areas of practical applicability gained by the research and experiments:

- Audit, and the professional fields performing an IT audit, can apply the procedures presented in the dissertation, which, among other things, contributes to more secure IT operation, compliance with regulatory requirements, and can increase the confidence of the organization's investors;
- Automated detection of control deficiencies can give a more objective picture of the quality of control;
- By improving the information security environment, organizations can be less exposed to external attackers and to intentional harm caused by employees, which can increase customer confidence in the organization's products / services, thus reducing reputational risks;
- Clarification of audit findings can reveal fraud, bottlenecks, and mitigate the risks of compromising the confidentiality, integrity, and availability of corporate data assets;
- The algorithm searching for the genetic potential of the described machine learning systems can be generalized, it works not only in a context-dependent way, but can also be used to find solutions to other business and professional problems processed by machine learning;
- The search procedure aimed at optimizing the size / composition of the training set, in case of a given target variable, it is able to select a more ideal direction based on the shifts of independent variables related to the target variable / for stochastically operating input-output systems;
- The mathematical derivation of model preference and the search for the ideal model can be similarly generalized, allowing the context-independent use of the algorithm where the hermeneutical subsystem has so far relied on human intuition to aggregate multiple layers of evaluation - so automation guarantees optimization and adherence to Knuth principles;
- Automate direction preferences to reduce the risk of human error;

During the research, a number of tasks and potential research directions were formulated in me, which I wanted to perform and document, but due to time and content constraints, it was not possible. The following points summarize the proposed research directions with which the results reported, can be improved:

- The research work has identified the detection of control deficiencies as a classification problem, however, regression algorithms may provide an opportunity to approximate the number of findings for controls;
- To estimate control gaps due to the confidentiality, nature of the business problem, a number of attributes were not available that could refine the predictions. If possible, it is recommended to repeat the experiment, where the sales revenue, information security budget, organizational hierarchy, size, etc. of each organization are available. These would presumably improve the judgment of the systems;
- The research did not separate the control deficiencies measured at the level of design, implementation and operational effectiveness related to each audit test, therefore the decision support system may be able to fine-tune this information to judge whether control deficiencies are already present in the control design, implementation or operation phases;
- A structured database with numerical values has been processed in the research, however, the textual reports may contain additional information that may help the accuracy of the predictions, so it is recommended to conduct an experiment combining the advantages of natural language processing with the methods described in the dissertation;
- The research did not attempt to find the perfect configuration, it was not intended due to content limitations. It is recommended to evaluate the algorithms after the optimal configuration. Furthermore, the search procedure described in the dissertation proves to be suitable for finding configurations more precisely;
- To increase the performance of the similarity analysis, it is recommended to develop the algorithm with universal approximators e.g. by neural network. Such improvements in anti-discriminatory models may make them suitable for refining deviation from the norm;
- In order to objectively determine direction preferences, it is recommended to experiment with additional algorithms that can provide more ideal results in terms of direction preferences, even in a hybrid solution that can reveal “lying” models and functions;

6. APPENDIX

6.1. References

1. BARTA, G. (2018a): Predicting Human Resource Attrition with Artificial Neural Networks. 55-66. p. In: ALMÁDI B. – GARAI-FODOR M. – SZEMERE P. T. (Szerk.): *Business as usual: Comparative socio-economic studies*. Budapest: Vízkapu Kiadó, 127 p.

2. BARTA, G. (2018b): The Increasing Role of IT Auditors in Financial Audit: Risks and Intelligent Answers. In: *Business Management and Education*, 16 (1) 81-93. p.

3. BARTA, G. – GÖRCSI, G. (2021): Risk Management Considerations for Artificial Intelligence Business Applications. In: *International Journal of Economics and Business research*, 21 (1) 87-106. p.

4. BARTA, G. – PITLIK, L. (2018): Startup felvásárlások multikulturális hátterének elemzése, avagy mesterséges intelligencia alapú ellenőrzőszámítás diszkriminancia-elemzéshez. 15-37. p. In: FARKAS A. (Szerk.): *A gazdaság kulturális szerkezete*. Gödöllő: Szent István Egyetemi Kiadó, 240 p.

5. BÁNKUTI, GY. (2010): About the method of component-based object comparison for objectivity. In: INTERNATION CONGRESS OF MATHEMATICIANS. (2010)(Hindustan). International Congress of Mathematicians: Abstracts, short communications, posters. Hindustan, Book Agency. p. 593-594.

6. BODA, M. A. (2019): Üzleti szimulációk és tanuló-rendszerek döntéshozatali mechanizmusai. Doktori disszertáció. Gödöllő: Szent István Egyetem, Gazdálkodás és Szervezéstudományok Doktori Iskola. 214 p.

7. FRIEDMAN, J. H. (2001): Greedy function approximation: A gradient boosting machine. In: *The Annals of Statistics*, 29 (5) 1189-1232. p.

8. GREENHOUSE L. (2005): Justices Unanimously Overturn Conviction of Arthur Andersen. https://www.nytimes.com/2005/05/31/business/justices-unanimously-overturn-conviction-of-arthur-andersen.html Keresőprogram: Google. Kulcsszavak: Arthur Andersen. Lekérdezés időpontja: 2020. 08. 08.

9. KÁSA, R. (2018): Neurális hálók alkalmazásának lehetőségei innovációs teljesítmény mérésére. In: *LOGISZTIKA - INFORMATIKA – MENEDZSMENT*, 3 (1) 60-73. p.

10. KNUTH, D. (1995): A=B. Előszó PETKOVSEK, M. – WILF, S. H. – ZEILBERGER, D. könyvében. Massachusetts: A K Peters/CRC Press. 224 p.

11. Magyar Nemzeti Bank (2015): A Jegybank azonnali hatállyal felfüggesztette a Buda-Cash Brókerház működési engedélyét és felügyeleti biztosokat rendelt ki. https://www.mnb.hu/sajtoszoba/sajtokozlemenyek/2015-evi-sajtokozlemenyek/a-jegybank-azonnali-hatallyal-felfuggesztette-a-buda-cash-broker haz-mukodesi-engedelyet-es-felugyeleti-biztosokat-rendelt-ki Keresőprogram: Google. Kulcsszavak: MNB sajtószoba, Buda-Cash Brókerház. Lekérdezés időpontja: 2020. 05. 02.

12. MASON, L. – BAXTER, J. – BARTLETT, P. – FREAN, M. (1999): Boosting algorithms as gradient descent. In: INTERNATIONAL CONFERENCE ON NEURAL INFORMATION PROCESSING SYSTEMS (12.)(1999)(Cambridge). NIPS'99: Proceedings of the 12th International Conference on Neural Information Processing Systems. Cambridge, p. 512-518.

13. PITLIK, L. – RIKK, J. – GÁNGÓ, V. – TÓTH, CS. (2020): A távoktatás, mint kritikus oktatási üzem – IT-aspektusai, avagy felkészülés a duális képzésre. In: *Magyar Internetes Agrárinformatikai Újság*, 23 (266) 1-26. p.

14. PITLIK, L. – VARGA, Z. – BARTA, G. – LOSONCZI, GY. – PITLIK, L. (jun.) – PITLIK, M. – PITLIK, M. (2017): Magyar statisztikai régiók érintettségi sorrendje a szálláshelyek árbevételének havi adatai alapján eltérő módszertanokkal. In: VIDÉKFEJLESZTÉSI KONFERENCIA (1.)(2017)(Szarvas). Magyar vidék - perspektívák, megoldások a XXI. században. Szarvas, Szent István Egyetem Egyetemi Kiadó. p. 127-140.

15. POMPON, R. (2016): IT Security Risk Control Management. An Audit Preparation Plan. Washington: Apress. 311 p.

16. STEGMAN E. – GUEVARA, J. – MICHELOGIKANNAKIS, N. – FUTELA, S. – SHARMA, S. – KAUSHAL, S. (2019): IT Key Metrics Data 2020: Industry Measures — Executive Summary. https://www.gartner.com/document/3975995?ref=gfeed Keresőprogram: Google. Kulcsszavak: IT Key metrics. Lekérdezés időpontja: 2020. 08. 02.

17. VASVÁRI, GY. (2008): Vállalati (szervezeti) kockázatmenedzsment. Budapest: Információs Társadalomért Alapítvány. 183 p.

18. VINOGRADOV, SZ. (2020): A nemzeti versenyképesség puha tényezői, a társadalmi versenyképesség. 109-138. p. In: CSATH M. (Szerk.):

Versenyképesség: új elméleti és módszertani megközelítések. Budapest: Dialóg Campus Kiadó, 215 p.

Standard

19. ISO/IEC 27001 (2013): Information technology – Security techniques – Information security management systems – Requirements. International Standard. 23 p.

6.2. List of publications in relation with the dissertation

Scientific Journal articles in Hungarian

1. BARTA, G. (2020): Tanúsítványok értékelése ellátási láncok IT biztonsági megfelelésének vizsgálatára. In: *Logisztikai trendek és legjobb gyakorlatok*, 6 (1) 27-30. p.

2. GÖRCSI, G. – SZÉLES, ZS. – BARTA, G. (2019): Üzleti intelligencia megoldások alkalmazásának sikertényezői - A hazai szolgáltató szektor nagyvállalatainak körében végzett mélyinterjús kutatás. In: *Információs Társadalom: Társadalomtudományi Folyóirat*, 19 (2) 23-34. p.

Scientific Journal articles in English

3. BARTA, G. (2018): Implementing and Evaluating Different Machine Learning Algorithms to Predict User Localization by the Strength of User Devices' Wi-Fi Signal. In: *SEFBIS Journal*, (12) 2-11. p.

4. BARTA, G. (2018): The Increasing Role of IT Auditors in Financial Audit: Risks and Intelligent Answers. In: *Business Management and Education*, 16 (1) 81-93. p.

5. BARTA, G. – GÖRCSI, G. (2021): Risk Management Considerations for Artificial Intelligence Business Applications. In: *International Journal of Economics and Business research*, 21 (1) 87-106. p.

Scientific other journal articles in Hungarian

6. BARTA, G. (2017): A Mesterséges Intelligencia hatása az üzleti folyamatokra. In: *Magyar Internetes Agrárinformatikai Újság*, 20 (233) 1-15. p.

7. BARTA, G. – PITLIK, L. (2020): Hipotézis-tervezés PhD-disszertációkhoz - Konzisztens gépi tanuló modellezés beltéri felhasználói lokalizáció meghatározásának pontosítására. In: *Magyar Internetes Agrárinformatikai Újság*, 23 (263) 1-13. p.

8. BARTA, G. – PITLIK, L. (2018): A Titanic katasztrófa túlélőinek becslése döntési fa alapú gépi tanuló eljárással. In: *Magyar Internetes Agrárinformatikai Újság*, 21 (234) 1-24. p.

Publications in conference proceedings in Hungarian

9. BARTA, G. – GÖRCSI, G. (2019): Csevegőrobotok a vállalati működésben. In: GAZDÁLKODÁS ÉS MENEDZSMENT TUDOMÁNYOS KONFERENCIA (3.)(2019)(Kecskemét). Versenyképesség és innováció. Kecskemét, Neumann János Egyetem Kertészeti és Vidékfejlesztési Kar. p. 912-917.

10. GÖRCSI, G. – BARTA, G. (2019): Az információs rendszer szerepe a döntési folyamatban. In: GAZDÁLKODÁS ÉS MENEDZSMENT TUDOMÁNYOS KONFERENCIA (3.)(2019)(Kecskemét). Versenyképesség és innováció. Kecskemét, Neumann János Egyetem Kertészeti és Vidékfejlesztési Kar. p. 252-256.

11. GÖRCSI, G. – BARTA, G. (2018): A CRM rendszerek szerepe a vevőkapcsolatok stratégiai kezelésében, vevőszegmentációs döntésekben. In: In: KÖZGAZDÁSZ DOKTORANDUSZOK ÉS KUTATÓK TÉLI KONFERENCIÁJA (4.)(2018)(Gödöllő). Közgazdász Doktoranduszok és Kutatók IV. Téli Konferenciája: Konferenciakötet. Budapest: Doktoranduszok Országos Szövetsége. p. 26-33.

12. PITLIK, L. – VARGA, Z. – BARTA, G. – LOSONCZI, GY. – PITLIK, L. (jun.) – PITLIK, M. – PITLIK, M. (2017): Magyar statisztikai régiók érintettségi sorrendje a szálláshelyek árbevételének havi adatai alapján eltérő módszertanokkal. In: VIDÉKFEJLESZTÉSI KONFERENCIA (1.)(2017)(Szarvas). Magyar vidék - perspektívák, megoldások a XXI. században. Szarvas, Szent István Egyetem Egyetemi Kiadó. p. 127-140.

Publications in conference proceedings in English

13. BARTA, G. (2018): Artificial Intelligence: Blessing or Curse? In: BUSINESS AND MANAGEMENT SCIENCES: NEW CHALLANGES IN THEORY AND PRACTICE (2018)(Gödöllő). *Proceedings of the International Conference "Business and Management Sciences: New Challenges in Theory and Practice"*. Volume I. Gödöllő, p. 141-145.

14. BARTA, G. (2018): Challenges in the compliance with the General Data Protection Regulation: Anonymization of personal information and related information security concerns. In: INTERNATIONAL SCIENTIFIC CONFERENCES OF THE FACULTY OF MANAGEMENT (10.)(2018)(Krakkó). Knowledge – Economy – Society. Business, Finance and Technology as Protection and Support for Society: proceedings. Krakkó, Cracow University of Economics. p. 115-121.

15. BARTA, G. – GÖRCSI, G. (2020): Assessing and managing business risks for artificial intelligence based business process automation. In: SCIENTIFIC CONFERENCE ON CONTEMPORARY ISSUES IN BUSINESS, MANAGEMENT AND ECONOMIC ENGINEERING (6.)(2019)(Vilnius). 6th Internation Scientific Conference Contemporary Issues in Business, Management and Economics Engineering '2019: proceedings. Vilnius, Vilnius Gediminas Technical University Press. p. 823-832.

16. BARTA, G. – GÖRCSI, G. (2018): Artificial Intelligence and Audit: Why is it necessary to audit the intelligent decision support? In: KÖZGAZDÁSZ DOKTORANDUSZOK ÉS KUTATÓK TÉLI KONFERENCIÁJA (4.)(2018)(Gödöllő). Közgazdász Doktoranduszok és Kutatók IV. Téli Konferenciája: Konferenciakötet. Budapest: Doktoranduszok Országos Szövetsége. p. 225-234.

17. BARTA, G. – GÖRCSI, G. (2017): Intelligent Decision Making and Process Automation for Public Organizations. In: INTERNATIONAL SCIENTIFIC CORRESPONDENCE CONFERENCE (5.)(2017)(Nitra). Legal, economic, managerial and environmental aspects of performance competencies by local authorities. Nitra, Slovak University of Agriculture in Nitra. p. 30-37.

18. BARTA, G. – LUDVAI, N. – PUSKÁS, A. (2020): The analysis of data privacy incidents and sanctions in Europe after GDPR enforcement. In: INTERNATIONAL WINTER CONFERENCE OF ECONOMICS PHD STUDENTS AND RESEARCHERS (6.)(2020)(Gödöllő). VI.

International Winter Conference of Economics PhD Students and Researchers: Conference Proceedings. Budapest, Association of Hungarian PhD and DLA Students. p. 35-48.

Book chapters

19. BARTA, G. (2018): Predicting Human Resource Attrition with Artificial Neural Networks. 55-66. p. In: ALMÁDI B. – GARAI-FODOR M. – SZEMERE P. T. (Szerk.): *Business as usual: Comparative socio-economic studies*. Budapest: Vízkapu Kiadó, 127 p.

20. BARTA, G. – PITLIK, L. (2018): Startup felvásárlások multikulturális hátterének elemzése, avagy mesterséges intelligencia alapú ellenőrzőszámítás diszkriminancia-elemzéshez. 15-37. p. In: FARKAS A. (Szerk.): *A gazdaság kulturális szerkezete*. Gödöllő: Szent István Egyetemi Kiadó, 240 p.

6.3. List of other publications

21. BARTA, G. (2015): Establishment of Enterprise Information Management Capability. In: INTERNATIONAL SCIENTIFIC CONFERENCES OF THE FACULTY OF MANAGEMENT (7.)(2015)(Krakkó). Knowledge-Economy-Society: Challenges for enterprises in knowledge-based economy. Krakkó, Cracow University of Economics. 29-36. p.

22. BARTA, G. – KARDOS, T. (2017): Integrált szervezeti rendszer bemutatása az Exxonmobil példáján keresztül. 49-56. p. In: SALAMONNÉ H. A. et al. (Szerk.): Ellátásilánc-menedzsment hallgatói esettanulmánykötet: Játszmák és stratégiai megoldások az ellátási lánc hálózatban. Budapest: Magyar Logisztikai Egyesület, 136 p.

23. BARTA, G. – ŁĘTEK, M. (2015): Equality in the Labor Market in Cracow. In: ACADEMIC INTERDISCIPLINARY CONFERENCE ON MODERN ECONOMICS AND SOCIAL SCIENCES (1.)(2015)(Tbilisi). International Academic Interdisciplinary Conference on Modern Economics and Social Sciences: proceedings. Tbilisi, Universal Publishing House. 281 p.

24. BARTA, G. – ŁĘTEK, M. (2015): Non-financial Performance Indicators as a New Approach to Measure the Value of Companies. In: INTERNATIONAL SCIENTIFIC CONFERENCES OF THE FACULTY OF MANAGEMENT (7.)(2015)(Krakkó). Knowledge-Economy-Society: Reorientation of paradigms and concepts of management in the contemporary economy. Krakkó, Cracow University of Economics. 247-254. p.

Articles under publication

25. BARTA, G. (2021): Gépi tanuló rendszerek audit kihívásai. In: *Gazdaságinformatikai Kutatási és Oktatási Fórum*, (12).