

Explainable Artificial Intelligence (XAI) Supporting Public Administration Processes – On the Potential of XAI in Tax Audit Processes

Nijat Mehdiyev

German Research Center for Artificial Intelligence (DFKI), Universität Saarland

Constantin Houy

German Research Center for Artificial Intelligence (DFKI), Universität Saarland

Oliver Gutermuth

German Research Center for Artificial Intelligence (DFKI), Universität Saarland

Lea Mayer

German Research Center for Artificial Intelligence (DFKI), Universität Saarland

Peter Fettke

German Research Center for Artificial Intelligence (DFKI), Universität Saarland

Follow this and additional works at: <https://aisel.aisnet.org/wi2021>

Mehdiyev, Nijat; Houy, Constantin; Gutermuth, Oliver; Mayer, Lea; and Fettke, Peter, "Explainable Artificial Intelligence (XAI) Supporting Public Administration Processes – On the Potential of XAI in Tax Audit Processes" (2021). *Wirtschaftsinformatik 2021 Proceedings*. 5.
<https://aisel.aisnet.org/wi2021/SSmartCity/Track08/5>

This material is brought to you by the Wirtschaftsinformatik at AIS Electronic Library (AISeL). It has been accepted for inclusion in Wirtschaftsinformatik 2021 Proceedings by an authorized administrator of AIS Electronic Library (AISeL). For more information, please contact elibrary@aisnet.org.

Explainable Artificial Intelligence (XAI) Supporting Public Administration Processes – On the Potential of XAI in Tax Audit Processes

Nijat Mehdiyev, Constantin Houy, Oliver Gutermuth, Lea Mayer, Peter Fettke

Germany Research Center for Artificial Intelligence (DFKI) and Saarland University,
Saarbrücken, Germany

{nijat.mehdiyev, constantin.houy, oliver.gutermuth, lea.mayer, peter.fettke}@dfki.de

Abstract. Artificial Intelligence (AI) can offer significant potential for public administrations which – in Germany – are likely to face considerable skills shortages in the next few years. AI systems can especially support the automation of processes and thus disburden administrative staff. As transparency and fairness play a major role in administrative processes, explainable AI (XAI) approaches are expected to enable a proper usage of AI in public administration. In this article, we investigate the potential of XAI for the support of tax authority processes, especially the selection of tax audit target organizations. We illustrate relevant tax audit scenarios and present the potential of different XAI techniques which we currently develop in these scenarios. It shows that XAI can significantly support tax audit preparations resulting in more efficient processes and a better performance of tax authorities concerning their main responsibilities. A further contribution of this article lies in the exemplary application of XAI usage guidelines in the public administration context.

Keywords: Explainable Artificial Intelligence, XAI, Public Administration, Tax Audit

1 Introduction

Public administrations in Germany will likely face significant skills shortages in the coming years according to a current PwC study [1]. In order to keep up the needed public services and to provide them timely and in a good service quality, more and more public administrations consider taking advantage of the potential of artificial intelligence (AI) to support public administration processes [2]. However, while AI applications generally need to be reliable, trustworthy and need to provide good results, e.g. correct predictions of events, AI solutions for public administrations especially must be transparent, fair, and non-discriminatory in their results. In this context, it is important that the outcomes delivered by an AI system, e.g. using machine learning (ML), and the way how these results came into existence are explainable. It is, e.g. important that human users can understand and follow the results which have been

produced by an ML technique, even if the approach is a “black box”. Explainable AI (XAI) approaches aim at providing and supporting the needed transparency of delivered outcomes by AI systems. While the potential of XAI methods has been broadly acknowledged for industrial scenarios, their potential for the support of public administration processes has not been intensely investigated, so far, and only little research exists in this field.

This article aims at contributing to the current state of research with an in-depth illustration of XAI potential for supporting public administration processes with a special focus on the tasks and duties of the German tax authorities. The methodical approach of this work is based on literature analysis and an in-depth case examination. The *main contribution* of this article is twofold. First, we define and describe various innovative AI-based use-cases for tax audit processes especially by following the key propositions recommended in the conceptual framework proposed by [3] for developing XAI solutions. By using this XAI design and deployment guideline in the context of public administration we were able to define and introduce the motivation and objectives for using generated explanations, the expected outcome and the context of explanation situation, the target audience and their expectations, preferences and requirements which impact the choice and implementation of the relevant XAI methods in public administration. Second, this study illustrates the applicability of XAI approaches by presenting the prototypically developed explanation solutions by using the semi-synthetically generated data based on the relevant variables identified by [4] which investigated similar processes in Austria by using conventional rule-based models. Our proposed solutions are presumed to support the proper selection of target organizations for tax audit as well, but by generating explanations on top of the applied black-box models. For the examined *process prediction* problem, it was of interest to examine two different use-cases and consequently to investigate two families of post-hoc explanation methods which support various user groups in their specific decision making objectives. On one hand, we applied three widely accepted local post-hoc explanations such as Shapley Values [5-6], LIME [7] and Individual Conditional Expectation (ICE) plots [8] which facilitate the auditors in justifying each individual decision provided by the black-box model. On the other hand, the global post-hoc explanation approaches such as feature importance, model tree-based regression surrogate model and Partial Dependence Plots (PDP) [9] were adopted which enable the management level decision makers or process owners to make more strategic decisions such as enhancing the business processes [10].

The article structure is as follows: after this introduction, section 2 presents conceptual foundations and related work. Section 3 introduces the tax audit context and motivates the tax audit scenario on which we focus in our in-depth case analysis as well as the description of the XAI potential and the discussion including our developed solution approaches. Section 4 discusses the findings before section 5 concludes the paper.

2 Conceptual Foundations and Related Work

2.1 Artificial Intelligence and Public Administrations

AI pursues the goal of developing technical systems that can solve problems for which a human being requires intelligence. Investigating the potential of AI in public administrations has been a trending topic in recent years. In their landscaping analysis for examining and classifying the AI implementations in public services [11], Misuraca et al. identify more than 85 AI application projects in the chosen European countries. In his study [12], Etscheid develops a framework for evaluating the opportunities of public administration processes which can be partially or fully automated through AI techniques. In [2], Djeflal analyzes and describes interesting application opportunities of AI in public administration while focusing on legal issues and the necessary free space for experimenting to tap its full potential for the improvement of public services. In this context, first normative guidelines for AI in public administrations have been discussed from different perspectives [13]. In [14], Wirtz et al. propose an integrated AI governance framework after examining the AI challenges and different AI regulation approaches for public administration. In their proposed assessment framework [15], van Noordt and Misuraca discuss the effects of AI in the government context by examining drivers impacting the adoption of AI and by analyzing the need for organizational changes. Other related work illustrates further interesting potential of AI in public administration processes or in the public sector. e.g. in the following concrete use-cases [16–18]:

1. Tax audit scenarios, especially the selection of tax audit target organizations: this case will be examined in detail in section 3.
2. AI-based traffic engineering and traffic management systems (TMS): AI can be used in complex Smart City environments to manage traffic based on current situational data, such as available capacities, the utilization status of certain routes or current sensor values, e.g. concerning air pollution etc.
3. Social welfare procedures: AI techniques from the field of natural language processing (NLP) and image recognition can support administrative back office procedures in the context of social welfare, e.g. the social integration assistance procedure or the assessment of personal handicap situations. In this context application forms and provided documents, e.g. medical reports, can be automatically analyzed, and decisions can be automatically prepared and suggested to the administrative staff based on this analysis.
4. AI-based application assessment for monetary support, e.g. for small and medium-sized enterprises (SME) and businesses in times of the Covid-19 pandemic: in the context of the current pandemic situation, German SMEs had the possibility to apply for monetary support. In this context, a sheer flood of applications was filed and could only be handled by random assessments while most of the applications were approved without a proper assessment because of staff shortage in the administrations. AI could provide more intense examinations of all applications filed in such

situations, which would be politically necessary as many cases of fraud have been identified in the meanwhile.

In such scenarios when using ML-based AI approaches, the developed results, e.g. predictive statements, provided decision support or automated decision making typically remain non-transparent, as the underlying ML approaches mostly function like “black boxes” and often provide no explanation why a certain conclusion has been drawn or why a certain decision has been proposed. However, this can be a severe showstopper for AI in public administration in Germany, because of the rigorous requirements concerning the transparency and fairness of decisions made by public administrations. In this context, decisions made and communicated to applicants should be “bulletproof”, as a large amount of decisions made by public administrations in Germany are challenged in court. If public administrations want to use AI to gain reliable and “bulletproof” support, the explainability of AI results and thus reliable XAI approaches gain more and more importance, which will be discussed in more detail in the following section.

2.2 XAI and its Necessity in Public Administration

To utilize the benefits of AI in the public administration while mitigating its risks, it is crucial to ensure the trustworthiness of the underlying systems. According to ethical guidelines drafted by a high level expert group on AI set up by the European Commission, three aspects should be considered throughout the entire lifecycle of intelligent systems in harmony to achieve the trustworthy AI [19]. These systems should comply with all laws and regulations (lawful), follow ethical principles and values (ethical) and should be technically and societally robust. A recent study proposes different recommendations at three different level (team, organization and industry) for building reliable, safe and trustworthy human-centered AI systems [20]. Another study suggests that various categories of trustworthiness approaches for AI such as fairness, explainability, auditability and safety should be considered when designing and implementing AI-based solutions [21]. In this context, apart from ensuring consideration of various requirements such as diversity, non-discrimination and fairness, technical robustness, privacy and data governance, societal and environmental well-being, etc. it is also essential to make the underlying AI-based systems accountable, responsible and transparent.

XAI methods have recently reemerged as crucial technical approaches to address these issues by making black-box AI systems more comprehensible, interpretable, and accountable [22]. The history of the research on making intelligent systems reaches back to more than three decades, where the initial studies concentrated mainly on making expert systems explainable [23-24]. The recent proliferation of advanced black-box machine learning systems has triggered a new generation of explainable artificial intelligence [25]. An overview of taxonomies of modern explanation methods, required quality desiderata, various evaluation mechanisms, types and tools and other relevant aspects and dimensions can be found in [10], [22], [26-27]. In the light of recent studies on application of XAI methods in different research and application domains such as healthcare, transportation, security, production, finance etc., we can observe the

potential of explanations on top of underlying advanced machine learning approaches for enhancing the decision making processes [28–30]. Considering the practical implications of explanatory systems in operationalizing the data-driven artificial advice givers and the lack of relevant research in the public administration domain, this study is a first attempt to introduce an approach that defines what prerequisites to consider and how to develop XAI solutions by examining tax audit processes. Furthermore, this study introduces and visualizes the outcomes of prominent XAI approaches for tax audit cases for illustration purposes.

3 On the Potential of XAI in Tax Audit Processes

3.1 Motivation and Case Context

In general, the key task of tax authorities is to make sure that taxes are paid properly, e.g. regarding the correctness of paid amounts, the equal and fair treatment of every person and organization who must pay taxes. Hence, this relates to both companies and private individuals. However, performing tax audits in companies are a much bigger challenge for tax authorities, since the amount of relevant information needed to do a proper tax audit is larger and more complex in companies. Furthermore, tax audits in companies can generate considerable additional tax revenues for states, which makes this task particularly important for tax authorities [16].

In Germany, tax audits are regulated by law, especially by §§ 193-207 of the German Fiscal Code (Abgabenordnung, AO). Possible types of audits are, e.g. tax audits, wage tax audits, special turnover tax audits, the abbreviated tax audit as defined in § 203 AO, as well as other audits such as standard rate audits. According to § 194 AO tax audits serve for the determination of the taxable person's tax circumstances and can comprise one or more tax types and taxation periods. The admissibility of a tax audit is regulated in § 193 AO in which is determined who may be subject to a tax audit. Within the scope of corporate tax audits, e.g. trade taxes, income taxes, corporate taxes, sales taxes, and other tax charges are among the factors that may be considered.

The data used by tax authorities for tax audits usually stem from financial accounting, asset accounting and payroll accounting. While these data provide tax authorities with a wide range of information for comprehensive analyses, the scope and depth of this audit information also imply considerable efforts and expenses. Although these efforts can be reduced by using IT systems, a complete analysis of all eligible companies or organizations is still challenging. Therefore, and for capacity reasons, companies and organizations that shall be audited are selected instead of deeply examining every single case [16].

The German tax audit regulations (Betriebsprüfungsordnung, BPO) are the basis for tax audits of the federal states' fiscal authorities and the Federal Central Tax Office (§ 1 BPO). The determination and evaluation of the tax-relevant facts must take place to ensure the equality of taxation (§§ 85, 199 AO). In legislative terms, this addresses the accurate determination and assessment of tax-relevant facts and does not aim at additional taxes [31]. At first, companies are classified according to their size: large

companies, medium-sized companies, small companies, or micro-entities. The classification of size is determined by the Federal Ministry of Finance (§ 3 BPO). According to § 4 BPO, large companies are audited without any gaps, the audit period directly follows the preceding audit period. Organizations in other size classes are often audited at irregular intervals [32]. The empirical outcomes of a recent study which used the real world data for analyzing tax audit processes in Austria have revealed that a segmentation of the firms in terms of various criteria including their size has significant implications on the performance of the applied machine learning models for identifying suspicious firms [4]. Such a classification increases the predictive strength of the underlying local models which is an important prerequisite for generating plausible explanations for auditors to justify the appropriateness of recommendations delivered by the data-driven intelligent systems. The following additional criteria can e.g. also be included in the selection: random audits, audit requests at the instigation of corporate tax offices, evaluation of operating data, sector audits, indicative rate audits, risk groups or evaluation of control material [31]. An important objective of tax authorities is to make the audit as efficient as possible while considering insufficient staffing and the principle of equal burden. Furthermore, organizations can apply for an audit because there is e.g., a special interest in a timely tax audit due to the sale of a business. This must also be considered in the context of discretionary selection [33].

For the tax audit selection, typical indications in a company's data are analyzed in more detail. The more suspicious aspects, inconsistencies, or potential violations against existing rules are found in one case, the more likely an in-depth analysis will be suggested [16]. It should be noted that, in addition to rule-based approaches and purely statistical procedures, the selection of audit cases can also be based on the auditors' experience in combination with indications in the data. In addition, other exogenous events such as applications of external auditors can also influence the selection of target organizations to be tested [33].

3.2 Issues

The tax audit which is referred to as a conclusive and retrospective review of the individual cases with regard to certain types of taxes and taxation periods is an essential tool for the corresponding authorities to accomplish their task of assessing and collecting taxes in accordance with the law. According to the reports of the Federal Ministry of Finance in Germany, 13,525 auditors were employed nationwide in the tax audits of the federal states in 2018 who were able to audit 188,973 enterprises out of the 7,816,301 firms registered in the business register of the tax offices [34]. These numbers suggest the audit coverage was just 2.4% which resulted in additional tax claims of 13.9 Billion Euros. A further analysis suggests that 10.9 Billion Euros additional taxes are claimed from large enterprises for which the audit coverage was 21.6%. From more than 5.5 Million of small enterprises only 1.1% were audited. These numbers especially related to the personnel shortage reveal the necessity of intelligent decision support to overcome the complications in enhancing the underlying processes. It is very important to emphasize that the identification of firms for examination is not sample-based and the processes behind a solid audit are very knowledge-intensive. The

auditors are therefore obliged to analyze potential cases over several years and provide a thorough justification for their choices. This time-consuming and mentally exhausting process is also affected by the heuristics and biases that prevent rational decision making.

Recent advancements in various branches of AI, particularly regarding ML, offer possibilities to enable such data-driven decision making for the identification of relevant cases for tax audit. Adoption of relevant machine learning approaches facilitates to develop intelligent systems that are capable to automate the tax audit processes fully or act as decision support systems to enhance the processes which cannot be automated due to legal, economic and technical factors. Moreover, due to the variety, volume, and veracity of the underlying data in the tax audit use-case, more complicated ML approaches with non-linear processing mechanisms are expected to provide superior results compared to conventional statistical approaches. The recent AI democratization attempts that aim at the wider adoption in daily work practices and making the participation of interdisciplinary communities possible have reduced the barriers to the accessibility of such innovative and advanced machine learning frameworks, methods, and applications. Therefore, the complexity in the current use-case does not solely lie in the adoption of advanced AI methods for public administration scenarios, but rather in their operationalization by incorporating them into business processes. Nevertheless, due to their black-box character, such advanced ML techniques including deep learning methods, suffer especially from the inability to provide appropriate explanations for their actions. Consequently, the lack of mechanisms which are required to verify the validity and robustness of models and to justify individual model assessments complicates the process of establishing human trust in developed intelligent systems.

3.3. Solution: XAI-Driven Intelligent Decision Support Systems

3.3.1. Machine Learning for Tax Audits with Strong Predictive Capabilities

Like other public administration processes in the tax audit use-case there is a need for intelligent decision support systems that can alleviate inefficiencies regarding the personnel shortage and due to the exhaustive nature of the problem. This, in turn, requires in the first step a systematic approach to structure the relevant machine learning projects and productionize the AI-based intelligent systems by making them explainable. To carry out a thorough, consistent, diligent and comprehensive predictive analytics and development project, it is of utmost importance to follow well established scientific [35] or industrial frameworks [36]. According to these guidelines it is crucial to understand the AI project requirements and functionalities for the tax audit processes both from business and technical perspectives, collect and prepare the required data, choose the appropriate modelling approaches and deploy them robustly after an evaluation phase.

In [4], Setnicka has examined the applicability of predictive analytics for the identification of the tax audit cases in Austria which is equivalent to the purpose of this

study. His study has investigated a data-driven identification of cases for subsequent claims and determination of fraudulent cases with the purpose to evaluate the quality of case selection and the audit process in general. After performing segmentation in terms of different criteria such as branches or the total amount of claims, various features such as turnover before taxes, operating expenses, assets in the balance sheet, total liabilities, the results of previous inspections, internal and external personnel costs, the amount of input taxes, travel expenses, maintenance costs etc. are collected. Although this path-breaking study has already addressed various important aspects and issues related to the tax audit, it adopted a comprehensible but less performant machine learning model, decision trees. Considering the imbalanced structure of the data and the non-linear relationships among the input features in the tax audit use-case and our previous experience from other public administration projects, it is conceivable to suggest that more advanced ML approaches such as deep learning or ensemble methods would provide more precise results. However, to enable collaboration among these ML techniques and the users in the tax audit scenario it is important to develop explanation solutions.

3.3.2. A Guideline for Structuring XAI Scenarios and Activities in Tax Audit

It is important to grasp that XAI is not a monolithic concept and the adoption of universal one-fits-all solutions for public administration is not necessarily reasonable [3]. Indeed, the sufficiency and relevance of the explanations are determined by the properties of the decision making environment, including but not limited to the user characteristics, the objectives of the explanation mechanisms, the nature of the underlying processes, the credibility and quality of predictions generated by adopted artificial intelligent systems [37]. Therefore, for developing solutions to generate relevant explanations for public administration processes, it is important to identify and describe the elements of these various factors and examine their interdependencies. For this purpose, this study uses a conceptual framework proposed in [3] by describing the key elements for designing an XAI solution in the tax audit use-case. This approach can also be used as a guideline for carrying out XAI projects in other public administration use-cases.

According to this concept it is crucial first to examine the characteristics of the stakeholders. Auditors, managers, data subjects (the firms), supervisory/regulatory authorities and AI developers are some of the stakeholders in the underlying tax audit use-case. These users who have different interests in using the XAI systems can be further classified in terms of their AI background (knowledge engineers vs. domain experts) or regarding their previous experiences (novice users vs. experienced users). The studies on explanations for intelligent systems suggest that the novice users prefer terminological explanations and may benefit from learning using the insights delivered by systems, whereas more experienced users opt for verifying the model knowledge [38, 39]. These users demand explanations for different objectives, e.g. verification of a reasoning trace in the adopted AI methods, justification/ratification of the reliability in the generated system outcomes, debugging the underlying AI models for improving accuracy and computational efficiency, learning from the system especially in the

absence of domain knowledge, improving effectiveness or efficiency for making good decisions fast etc. By using these explanations tailored to their preferences, not only trust in the AI system is established but the users can also analyze the technical robustness of the model by checking its consistency, stability and representativeness or examine the reasons if the algorithmic fairness is violated. Various XAI approaches are proposed to address these requirements. These approaches can be model-specific which imply that they explain only particular models or model-agnostic which are independent of the adopted ML models. Furthermore, the scope of the explanations can be global or local depending whether a single outcome or the whole model behavior for the total population is explained. Finally, explanations can be generated before training models (pre-model), during the model training (in-model) and after obtaining the results from trained models (post-model).

After examining the characteristics of the process stakeholders and interdependencies among various process specific and decision making environment-related aspects, this study identifies various scenarios for the tax audit use-case for which different XAI solutions should be developed (see Figure 1 on the next page). We would like to note that the introduced four scenarios are chosen to demonstrate how diversified the explanation solutions can be within a single use-case, tax audit. The list of potential scenarios can be easily extended by addressing the requirements of different stakeholders, their multifaceted explanation objectives, or the context of the predictions. In the first scenario, the proposed XAI solution is assumed to facilitate the tax auditors to ratify the data-driven recommendations. These users are responsible for making decisions for identifying suspicious firms and have interest to justify individual model decisions. They pursue mainly the purpose whether the provided AI decisions regarding the proposed audit cases are reliable and reasonable. Thus, for this audience it is appropriate to develop local post-hoc explanation approaches such as ICE Plots, Shapley values, LIME or other alternative local surrogate models, case-based explanations, counterfactual explanations etc.

On another side, the management is concerned with making strategic decisions to enhance the underlying processes by identifying process bottlenecks in relation to the identification of the cases for audit and developing measures to boost the effectiveness and efficiency of case handling by understanding the implications of the underlying AI models. Thus for this XAI scenario 2, the global post-hoc explanation approaches such as PDP, global surrogate models, Shapley dependent plots, Shapley Summary Plots, Accumulated Local Effects (ALE) Plots etc. are more suitable since they allow the decision-makers to examine all decisions made by the model globally [40]. In the third scenario, the supervisory/regulatory authorities may have an interest in carefully identifying samples for testing or verifying the availability of discrimination elements. Since they also lack in-depth technical AI knowledge like the first two stakeholders, post-hoc explanations are also relevant here. The scope of interpretability depends on their objective of examination. Finally, the AI developers are more interested in validating the inner working mechanism and reasoning trace of black-box AI approaches. By using e.g., the explanations generated by intrinsic model-specific approaches, they can debug models by understanding the problematic areas.

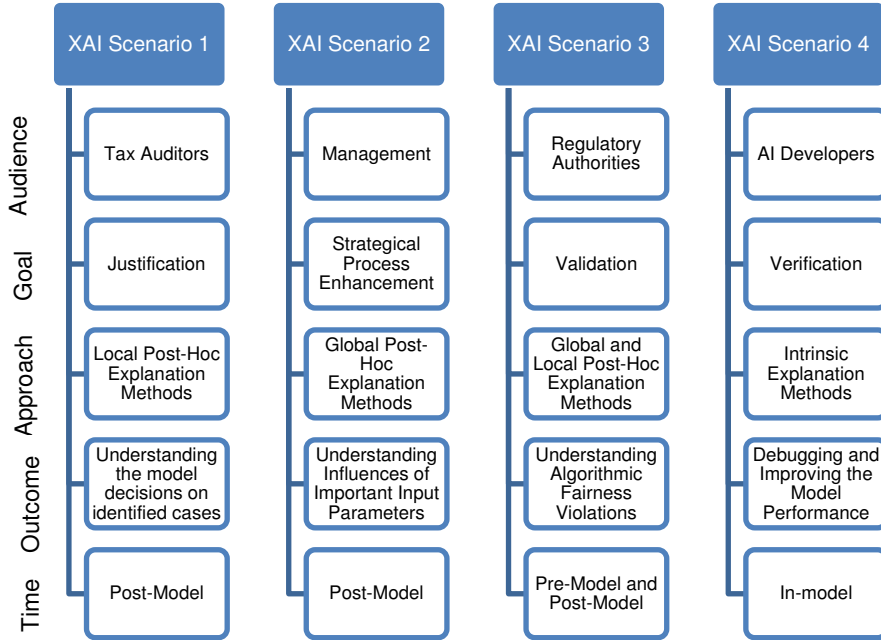


Figure 1. XAI Scenarios for Tax Audit

3.3.3. Demonstration of Chosen XAI Methods for Tax Audit

For illustrative purposes, we have implemented various local and global post-hoc explanation approaches on top of the applied gradient boosting machine (GBM) model for the binary classification problem and visualized the obtained explanations by using synthetic data based on the variables identified by [4]. The main idea of ICE is examining the influence of individual variables by analyzing each feature separately (see Figure 2). After choosing the variable of interest, the new data instances are created by making marginal changes to it through the identified grid of values and by keeping the values of other variables constant for the examined instance. The created data instances are scored by the black-box model and prediction scores on these value changes are then visualized.

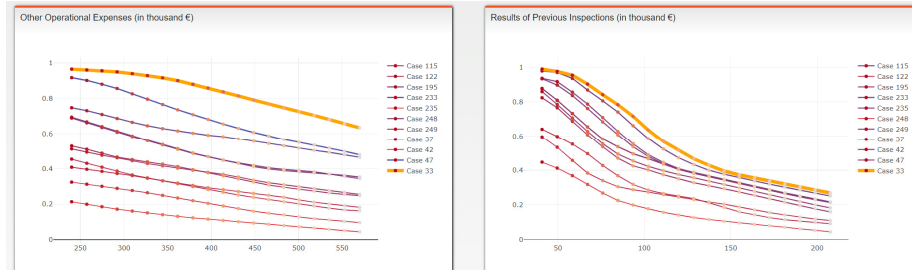


Figure 2. Individual Conditional Expectation (ICE) Plots

The steps of LIME, a perturbation-based local explanation approach, include perturbation of the dataset for the chosen instance, the scoring by the adopted black-box approach, weighting the generated new instances by calculating the distances to the original examined instance, fitting a comprehensible technique by using the original input values and prediction scores and presenting the extracted feature weights as explanations (see Figure 3).

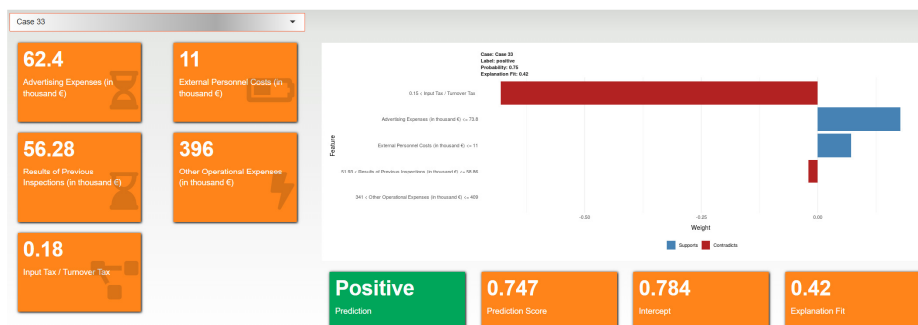


Figure 3. LIME Explanations

Shapley values is another feature attribution-based explanation method which stems from cooperative game theory. For the machine learning explanation situation, the input variable values of the examined observations are considered as cooperative game players and the model predictions are the respective payoffs. The sum of the obtained Shapley values illustrates and explains the contribution of each variable to the deviation between obtained prediction for the observation and average model prediction for the examined dataset (see Figure 4).

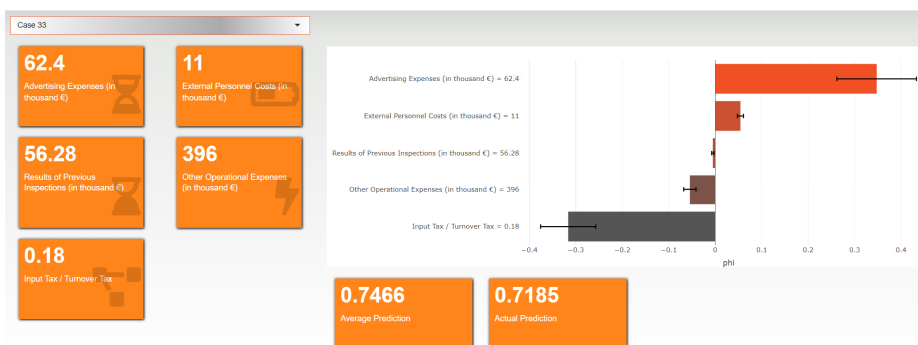


Figure 4. Shapley Values

Figure 5, 6 and 7 present the global post-hoc explanations from our dashboard namely, the surrogate regression model tree, PDP and the model-specific feature importance graph respectively. The main objective of global surrogate models is making the black-box model understandable by approximating its decisions. In our use-

case we score the validation dataset by using the adopted black-box model, gradient boosting machine. Once the prediction scores are available, a white-box model of choice is fitted by using the original feature scores values as input data and prediction scores as output data. An overview of the comprehensible models that can be used as surrogate models can be found in [41]. In our use-case we implemented M5 regression model trees that combine the decision trees and linear models. The learned representations are then used as explanations (see Figure 5).

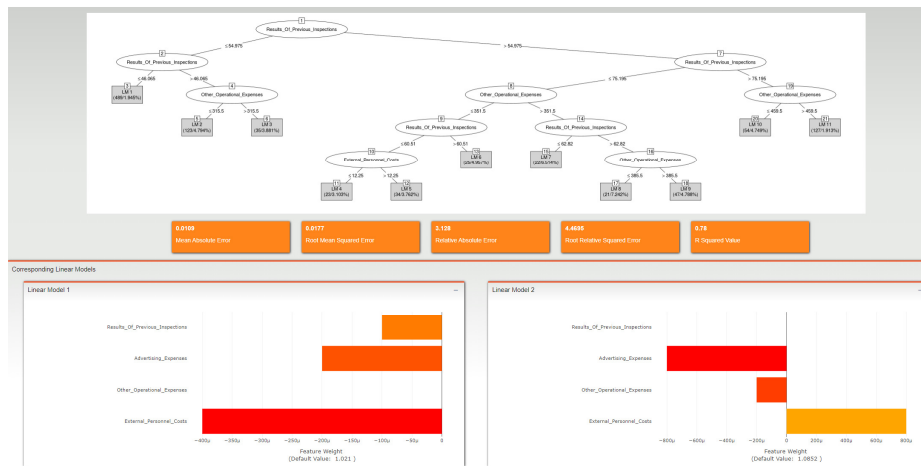


Figure 5. Global Surrogate Regression Model Tree

PDP is a widely recognized model-agnostic approach that examines the effects of marginal changes of chosen variables to the response of the black-box predictors (see Figure 6). It can be seen as the average of ICE plots described earlier; however, it provides a global interpretation. In [42], Zhao and Hastie suggest that under specific circumstances the PDP can be used to generate causal explanations.

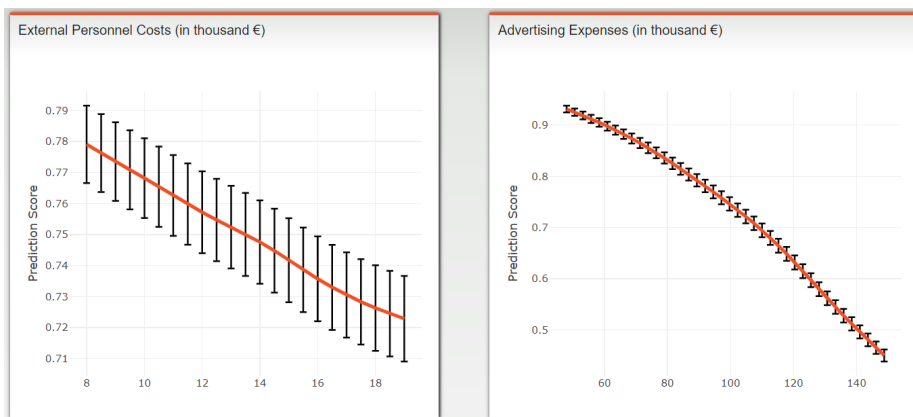


Figure 6. Partial Dependence Plots (PDP)

Finally, the feature importance graph reveals what features are seen as important by the model when making the decisions. In our study, we use the model-specific approach for neural networks (see Figure 7). Alternatively, the model agnostic approaches can be adopted. The practical applications and findings also suggest that using the explanation generated by different XAI methods together can generate more value since it allows the users to examine the model outcomes from different perspectives. For this purpose, it is essential to ensure harmonic transitions among different explanations by avoiding the confusions for users.

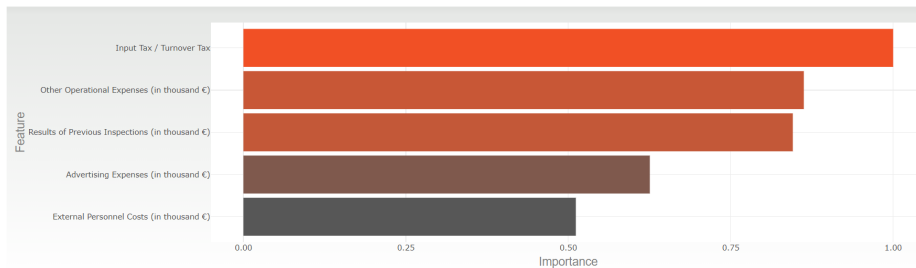


Figure 7. Feature Importance

4 Discussion

Public sector can benefit from AI projects that are expected to enhance the public good [43]. At the same time, such projects should not restrict themselves to increasing the efficiency or reducing costs only but generate added value for different stakeholders with various, sometimes conflicting objectives [44]. Various requirements such as transparency, safety, reliability, trustworthiness and fairness for designing, developing and deploying intelligent systems for supporting the decision making processes pose challenges by increasing the complexity of AI projects in public administration [45]. Furthermore, in difference to private sector projects, public sector projects often face scrutiny and oversight which increases the demand for explainability [45].

A well performed machine learning model and generated relevant explanations play a crucial role for establishing a data-driven culture in the public administration which is an essential aspect of a digital transformation. By bringing the people, processes, and technology together, this culture has great implications for enhancing the processes and improving “business models” in public administration that result in great advancements in provided services. Therefore, it is crucial to reduce the barriers between human users and data-driven intelligent systems and enable the human-in-the-loop concept. In this regard, this study aims to explore the potential of XAI for decision making processes in public administration particularly by examining the decision support structures for tax audit processes. Considering the research gap regarding the development and application of the explanation methods and corresponding interfaces in the public administration context, this study is one of the first attempts to present an approach to structuring the relevant XAI activities in this domain. Since the requirements of the users and the context of the explanation situation significantly influence the

appropriateness of the generated explanations, such a holistic approach is required to mitigate the risks that can lead to the prospective failures.

Generating explanations for black-box machine learning methods with high quality and conveying them to the relevant stakeholders in a suitable and robust manner also facilitates creating added value by productionizing the intelligent systems. Some of the benefits offered by explanations for the tax audit use-case are (i) effectiveness: the explanations enable the auditors to identify the potential audit cases more precisely by making good decisions, (ii) efficiency: the explanations facilitate the auditors to make decision more quickly, (iii) persuasiveness: considering the bureaucratic and legal complexities it is hard to convince the users to use machine learning decisions and relevant explanations can help to overcome this challenge, (iv) education: by using consistent and robust explanations especially the novel users can learn from the intelligent system about the relationships among various factors, (v) trust: explanations are sought to increase the auditors trust in the artificial advice givers [46].

For illustration purposes the chosen prominent XAI approaches from global and local post-hoc explanation families are presented in this study as well by discussing their suitability for various use-cases in tax audit domain. However, it is important to note that these solutions were developed based on the semi-synthetically generated data and have mostly informative character for the corresponding research community. In the future work, we aim to develop the XAI solutions by examining the processes with corresponding authorities, to integrate the relevant explanation approaches by considering various desiderata for explanations and finally to perform a thorough evaluation by using the relevant measures and mechanisms.

5 Conclusion

In this paper we have examined the potential which XAI can offer for supporting public administration processes with a focus on the tasks and duties of German tax authorities. We concentrated on the process of preparing tax audits by supporting advantageous selections of target organizations for tax audits. We have, furthermore, presented according XAI solutions which we are currently developing, e.g. the XAI-based information dashboard supporting the selection of target organizations by providing an in-depth explanation of the major influencing factors which lead to the selection suggestions of the AI system. Responsibility, fairness, transparency, and the avoidance of discrimination are some of the key values of public administration. XAI methods can significantly contribute to the realization and implementation of these key values when using AI to support public administration processes. We have demonstrated this in our in-depth case description. However, the current efforts for further developing these approaches are definitely needed to provide a better acceptance and trust into AI systems, which is particularly necessary in the context of public administrations.

References

1. Detemple, P., Höhn, A.: Fachkräftemangel im öffentlichen Dienst. Prognose und Handlungsstrategien bis 2030. Stud. der PricewaterhouseCoopers (2018).
2. Djeflal, C.: Künstliche Intelligenz in der öffentlichen Verwaltung (Artificial Intelligence in Public Administration). Berichte des NEGZ. 3 (2018).
3. Mehdiyev, N., Fettke, P.: Explainable Artificial Intelligence for Process Mining: A General Overview and Application of a Novel Local Explanation Approach for Predictive Process Monitoring. arXiv Prepr. arXiv2009.02098 (2020).
4. Setnicka, M.: Predictive Analytics in der österreichischen Finanzverwaltung. Informatik (2016).
5. Štrumbelj, E., Kononenko, I.: Explaining prediction models and individual predictions with feature contributions. Knowl Inf Syst. 41, 647–665 (2014).
6. Lundberg, S.M., Lee, S.-I.: A unified approach to interpreting model predictions. In: 31st Conference on Neural Information Processing Systems, pp. 4765–4774. Long Beach, CA, USA (2017).
7. Ribeiro, M.T., Singh, S., Guestrin, C.: Why should I trust you? Explaining the predictions of any classifier. In: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pp. 1135–1144. ACM Press, New York (2016).
8. Goldstein, A., Kapelner, A., Bleich, J., Pitkin, E.: Peeking inside the black box: Visualizing statistical learning with plots of individual conditional expectation. J. Comput. Graph. Stat. 24, 44–65 (2015).
9. Friedman, J.H.: Greedy function approximation: a gradient boosting machine. Ann. Stat., 1189–1232 (2001).
10. Lipton, Z.C.: The mythos of model interpretability. Queue 16, 31–57 (2018).
11. Misuraca, G., van Noordt, C., Boukli, A.: The use of AI in public services: results from a preliminary mapping across the EU. In: Proceedings of the 13th International Conference on Theory and Practice of Electronic Governance, pp. 90–99. ACM Press, (2020).
12. Etscheid, J.: Artificial Intelligence in Public Administration. In: International Conference on Electronic Government. pp. 248–261. Springer (2019).
13. Djeflal, C.: Artificial Intelligence and Public Governance: Normative Guidelines for Artificial Intelligence in Government and Public Administration. In: Wischmeyer, T., Rademacher T. (eds.) Regulating Artificial Intelligence. pp. 277–293. Springer, Switzerland (2020).
14. Wirtz, B.W., Weyerer, J.C., Sturm, B.J.: The Dark Sides of Artificial Intelligence: An Integrated AI Governance Framework for Public Administration. Int. J. Public Adm. 43, 818–829 (2020).
15. van Noordt, C., Misuraca, G.: Evaluating the impact of artificial intelligence technologies in public services: towards an assessment framework. In: Proceedings of the 13th International Conference on Theory and Practice of Electronic Governance, pp. 8–16. Athens 2020.
16. Fettke, P.: Digitale Betriebsprüfung. In: Klenk T., Nullmeier F., Wewer G. (eds) Handbuch Digitalisierung in Staat und Verwaltung, pp. 1-12. Springer VS, Wiesbaden

- (2019).
17. Bauer, W., Riedel, O., Braun, S., Etscheid, J., Von Lucke, J., Stroh, F.: Künstliche Intelligenz in der Öffentlichen Verwaltung - Anwendungsfelder und Szenarien. Fraunhofer-Institut für Arbeitswirtschaft und Organ (2020).
 18. Houy, C., Gutermuth, O., Fettke, P., Loos, P.: Potentiale künstlicher Intelligenz zur Unterstützung von Sachbearbeitungsprozessen im Sozialwesen. Berichte des NEGZ 8, 1–32 (2020).
 19. Ethics Guidelines for Trustworthy AI. Report, European Commission (2019).
 20. Shneiderman, B.: Bridging the Gap Between Ethics and Practice: Guidelines for Reliable, Safe, and Trustworthy Human-centered AI Systems. *ACM Trans. Interact. Intell. Syst.* 10, 1–31 (2020).
 21. Toreini, E., Aitken, M., Coopamootoo, K., Elliott, K., Zelaya, C.G., van Moorsel, A.: The relationship between trust in AI and trustworthy machine learning technologies. In: Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency, pp. 272–283. (2020).
 22. Doshi-Velez, F., Kim, B.: Towards A Rigorous Science of Interpretable Machine Learning. *arXiv Prepr. arXiv:1702.08608* (2017).
 23. Swartout, W.R., Moore, J.D.: Explanation in second generation expert systems. In: David JM., Krivine JP., Simmons R. (eds) *Second generation expert systems*. pp. 543–585. Springer, Berlin, Heidelberg (1993).
 24. Wick, M.R., Thompson, W.B.: Reconstructive expert system explanation. *Artif. Intell.* 54, 33–70 (1992).
 25. Gunning, D., Aha, D.W.: DARPA’s explainable artificial intelligence program. *AI Mag.* 40, 44–58 (2019).
 26. Guidotti, R., Monreale, A., Ruggieri, S., Turini, F., Pedreschi, D., Giannotti, F.: A Survey Of Methods For Explaining Black Box Models. *ACM Comput. Surv.* 51.5, 1–42 (2018).
 27. Miller, T.: Explanation in artificial intelligence: Insights from the social sciences. *Artif. Intell.* 267, 1–38 (2019).
 28. Tjoa, E., Guan, C.: A survey on explainable artificial intelligence (XAI): Towards medical XAI. *arXiv Prepr. arXiv:1907.07374*. (2019).
 29. Arrieta, A.B., Díaz-Rodríguez, N., Del Ser, J., Bennetot, A., Tabik, S., Barbado, A., García, S., Gil-López, S., Molina, D., Benjamins, R.: Explainable Artificial Intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI. *Inf. Fusion* 58, 82–115 (2020).
 30. Rehse, J.-R., Mehdiyev, N., Fettke, P.: Towards Explainable Process Predictions for Industry 4.0 in the DFKI-Smart-Lego-Factory. *KI-Künstliche Intelligenz* 33.2, 181–187 (2019).
 31. Buck, R., Klopfer, M.: *Betriebsprüfung*. Springer, Wiesbaden (2011).
 32. Waschbusch, G., Zieger, G.L.: Die steuerliche Betriebsprüfung als wichtigste Form der Außenprüfung - Grundlagen und Ablauf einer Betriebsprüfung (Teil I). *Der Steuerberater*, 329–225 (2016).
 33. Mösbauer, H.: *Steuerliche Außenprüfung: (Betriebsprüfung) - Steuerfahndung - Steueraufsicht*. Oldenbourg Verlag, München (2005).
 34. Bundesfinanzministerium: *Ergebnisse der steuerlichen Betriebsprüfung 2018*.

- Monatsbericht des BMF (2019).
35. Shmueli, G., Koppius, O.R.: Predictive Analytics in Information Systems. *MIS Q.* 35, 553–572 (2011).
 36. Chapman, P., Clinton, J., Kerber, R., Khabaza, T., Reinartz, T., Shearer, C., Wirth, R.: CRISP-DM 1.0: Step-by-step data mining guide. SPSS inc. 9, 13 (2000).
 37. Mehdiyev, N., Fettke, P.: Prescriptive Process Analytics with Deep Learning and Explainable Artificial Intelligence. In: European Conference on Information Systems (ECIS) (2020).
 38. Gregor, S., Benbasat, I.: Explanations from intelligent systems: theoretical foundations and implications for practice. *MIS Q.* 23, 497–530 (1999).
 39. Ji-Ye Mao, I.B.: The use of explanations in knowledge-based systems: Cognitive perspectives and a process-tracing analysis. *J. Manag. Inf. Syst.* 17, 153–179 (2000).
 40. Adadi, A., Berrada, M.: Peeking inside the black-box: A survey on Explainable Artificial Intelligence (XAI). *IEEE Access* 6, 52138–52160 (2018).
 41. Freitas, A.A.: Comprehensible classification models. *ACM SIGKDD Explor. Newsl.* 15.1, 1–10 (2014).
 42. Zhao, Q., Hastie, T.: Causal interpretations of black-box models. *J. Bus. Econ. Stat.*, 1–19 (2019).
 43. Cath, C., Wachter, S., Mittelstadt, B., Taddeo, M., Floridi, L.: Artificial intelligence and the ‘good society’: the US, EU, and UK approach. *Sci. Eng. Ethics.* 24.2, 505–528 (2018).
 44. Crawford, K.: Can an algorithm be agonistic? Ten scenes from life in calculated publics. *Sci. Technol. Hum. Values.* 41.1, 77–92 (2016).
 45. Desouza, K.C., Dawson, G.S., Chenok, D.: Designing, developing, and deploying artificial intelligence systems: Lessons from and for the public sector. *Bus. Horiz.* 63.2, 205–213 (2020).
 46. Nunes, I., Jannach, D.: A systematic review and taxonomy of explanations in decision support and recommender systems. *User Model. User-adapt. Interact.* 27-3-5, 393–444 (2017).