Applying Explainable Artificial Intelligence Techniques on Linked Open Government Data

Evangelos Kalampokis, Areti Karamanou, Konstantinos Tarabanis

Applying Explainable Artificial Intelligence Techniques on Linked Open Government Data

Evangelos Kalampokis[0000-0003-4416-8764], Areti Karamanou[0000-0003-2357-9169], Konstantinos Tarabanis[0000-0002-4663-2113]

Centre for Research & Technology Hellas (CERTH), Greece
University of Macedonia, Thessaloniki, Greece
{ekal,akarm,kat}@iti.gr

Abstract. Machine learning and artificial intelligence models have the potential to streamline public services and policy making. Frequently, however, the patterns a model uncovers can be more important than the model's performance. Explainable Artificial Intelligence (XAI) have been recently introduced as a set of techniques that enable explaining individual decisions made by a model. Although XAI has been proved important in various domains, the need of using relevant techniques in public administration has only recently emerged. The objective of this paper is to explore the value and the feasibility of creating XAI models using high quality open government data that are provided in the form of linked open statistical data. Towards this end, a process for exploiting linked open statistical data in the creation of explainable models is presented. Moreover, a case study where linked data from the Scottish open statistics portal is exploited in order to predict and interpret the probability the mean house price of a data zone to be higher than the average price in Scotland is described. The XGBoost algorithm is used to create the predictive model and the SHAP framework to explain it.

Keywords: Linked Data, Open Government Data, Machine Learning, Artificial Intelligence, XAI, SHAP, XGBoost.

1 Introduction

Machine learning and artificial intelligence models have been recently employed to improve public services and policy making [25]. In many cases these models have been proven accurate to predict the outcome of relevant events and thus effective to support public administration and policy makers [22], [12]. It is common, however, the patterns a model uncovers to be more important than the model's prediction performance [3]. Explainable Artificial Intelligence (XAI) techniques have been recently introduced to explain individual decisions made by the model [1]. XAI aims at producing more explainable models while maintaining a high level of learning performance (e.g., prediction accuracy), and enabling humans to understand, appropriately trust, and effectively manage the emerging generation of artificially intelligent partners [7]. XAI has been already proved particularly important in medical applications [13] and in transport [18]. However, the need for using relevant techniques in policy making and public administration only recently has emerged, and their value and feasibility are still not clear [8].

At the same time, Open Government Data (OGD) can play an important role in creating machine learning and artificial intelligence models [6]. Linked Data technologies can further contribute towards this direction because they enable the creation and dissemination of high-quality data that can be easily combined across disparate sources [10]. In particular, linked open statistical data provide statistics such as demographics (e.g., census data), economic, and social indicators (e.g., number of new businesses, unemployment). In statistics, linked data enable the application of analytics on top of disparate and previously isolated datasets.

The objective of this paper is to explore the value and the feasibility of creating XAI models using high quality OGD that are provided in the form of linked open statistical data. Towards this end, a process for exploiting linked open statistical data for the creation of explainable models is presented. Moreover, a case study where linked data from the Scottish open statistics portal are exploited in order to predict and interpret the probability the mean house price of a data zone to be higher than the average price in Scotland. eXtreme Gradient Boosting (XGBoost) is used to create the predictive model [2] and the SHapley Additive exPlanation (SHAP) framework to explain the predictive model [16].

The rest of this paper is structured as follows: In section 2 the process for applying XAI techniques on Linked Open Government Data is presented. Section 3 presents the research approach that is followed to achieve the objective of the paper. Section 4 describes the case of creating an explainable predictive model using data from the Scottish open data portal. Finally, section 5 summarizes the results and identifies open research issues.

2 Create Explainable Models using LOGD

By adapting explainable machine learning processes in the literature [21], the four broad steps of our approach are defined:

Specifying the problem: A supervised machine learning problem can be specified as either regression or classification. Government data and statistics typically include continuous variables and thus natively support regression analysis. However, classification problems can be addressed more easily and with higher accuracy. So, it is important in this step to transform a regression problem into a classification one. Moreover, the level of analysis should be decided in order to enable the use of a big number of data samples. For example, fine grained geographical areas (e.g., LAU levels in Europe) or time periods should be considered. Finally, the setting of the problem can be based either on time series analysis or on tabular analysis.

Collecting the data: Today, a large volume of statistical data is disseminated using linked data technologies [9]. This linked open statistical data is provided on the Web through official open government data portals launched by organizations and public authorities. Examples include the data portals of the Scottish and Japan's (e-Stat) governments, the data portal of the environmental department of the Flemish government (VLO), DCLG in the UK, and the data portals that host the Italian (ISTAT) and Irish (Irish CSO) 2011 censuses. The connection of these data portals would create a Knowledge Graph of quality and fine-grained statistical data including demographical, social and business indicators across countries. This knowledge graph, that would facilitate data discovery and collection, can be created by identifying [9] and addressing [11] interoperability challenges for connecting statistical data from multiple trustworthy sources. All the official portals provide SPARQL endpoints and thus the data can be easily collected by specifying and submitting relevant queries.

Creating the predictive model: In this step the actual predictive model is created. This includes selecting the algorithm, tuning the model to the optimal hyper-parameter values, and selecting the evaluation score. The fact that the result of a SPARQL query is a tabular-style dataset should be taken into account. For example, tree-based models can be more accurate than neural networks in many applications. While deep learning models are more appropriate in fields like image recognition, speech recognition, and natural language processing, tree-based models consistently outperform standard deep models on tabular-style datasets [2].

Explaining the predictive model: An explanation is the collection of features in the interpretable domain that have contributed to produce a decision (e.g., classification or regression) [17] for a given example. Various approaches have been proposed for explaining model predictions varying in scope and flexibility [19]. The scope indicates whether the method generates global explanations or local explanations, whereas the flexibility indicates whether the approach is model-specific or model-agnostic. Local explanations reveal the impact of input features on individual predictions of a single sample. Two recently proposed model-agnostic methods are the linear interpretable model-agnostic explainer (LIME) [20] and Shapley additive explanations (SHAP) [14]. Although SHAP is a local explainability model, it introduces global interpretation methods based on aggregations of Shapley values. Due to their generality, these methods have been used to explain a number of classifiers, such as neural networks and complex ensemble models, and in various domains ranging from law, medicine, finance, and science [24].

3 Research Approach

In order to demonstrate the applicability and value of applying XAI techniques on open government data, the Scottish government data portal (http://statistics.gov.scot) providing statistical data for free reuse is employed.

Currently, the portal contains 250 datasets covering various societal and business aspects of Scotland at different granularity levels. Data Zones refer to the primary geography for the release of small areas statistics in Scotland, while Council Areas are the coarser geographical units in Scotland.

The portal utilizes linked data technologies in order to improve data quality and also to make available the data as a unified knowledge graph. The different datasets are connected through typed links (mainly using the RDF Data Cube vocabulary) enabling users to search in a uniform way across all the available datasets and to easily combine data from different datasets. The portal has released a SPARQL endpoint (https://statistics.gov.scot/sparql), where users can submit queries to retrieve data.

In this case, eXtreme Gradient Boosting (XGBoost) is used to create the predictive model. XGBoost is an implementation of a generalized gradient boosting algorithm [2]. Boosting refers to the general problem of boosting the performance of weak learning algorithms by combining all the generated hypotheses into a single hypothesis [4]. This idea was further elaborated in gradient boosting [5]. In this case, one new weak learner is added at a time and existing weak learners in the model are frozen and left unchanged. The XGBoost algorithm has been applied to many domains, such as transportation, health, and energy, because of its high speed, high accuracy, and good robustness. It is indicative that during 2015, the 17 out of 29 winning solutions that were submitted to Kaggle competitions used XGBoost [2].

Moreover, the SHapley Additive exPlanation (SHAP) framework is employed for explaining the predictive model. SHAP, a local explainability model that is based on Shapley values, is employed [16]. The Shapley value method is a game theory method that assigns payouts to players depending on their contribution to the total payout where players cooperate in a coalition [23]. In machine learning the "game" is the prediction task for a single instance of the dataset. The "gain" is the prediction minus the average prediction of all instances and the "players" are the feature values of the instance that collaborate to receive the gain. The Shapley value is the average marginal contribution of a feature value across all possible coalitions.

The Shapley value method is computationally expensive because going over all coalitions scales exponentially with the increase in the number of features. SHAP solved this problem by enabling the exact computation of Shapley values in low order polynomial time instead of exponential by leveraging the internal structure of tree-based models [14], [15]. SHAP also proposed SHAP interaction values, which are an extension of Shapley values that directly capture pairwise interaction effects. Moreover, SHAP introduced global interpretation methods based on aggregations of Shapley values such as the SHAP feature importance, which is measured as the mean absolute Shapley values.

4 Predicting and Explaining House Prices in Scotland

In this section, the results of applying XAI techniques to linked open government data is presented according to the four broad steps defined in Section 2 and the detailed setting described in Section 3.

4.1 Specifying the Problem

In our case, the 2011 Data Zones are employed as the geographical units of analysis for the case study. Data zones are the core geography for dissemination of results from Scottish Neighborhood Statistics. They are designed to have roughly standard populations of 500 to 1,000 household residents. There are 6,976 2011 Data Zones.

The topic that will be explored is the mean house prices in the 2011 Data Zones in the year 2017. The average mean house price in 2017 across the 6,976 data zones is £168,285, while the median mean house price is £148,375. The problem that will be explored is the prediction of the probability the mean house price of a data zone is higher than the average price (£168,285). Moreover, what are the factors that contribute towards this prediction in each data zone.

4.2 Collecting Data from the Scottish LOGD Portal

In order to solve the specified problem, compatible datasets that can be exploited are collected from the Scottish open data portal. Towards this end, two criteria are specified: (a) the dataset includes data for 2017 as the year of reference, and (b) the dataset includes data at the granularity level of 2011 data zones. Because the available data have been shaped as Linked Data, a SPARQL query is structured in order to formally specify these criteria. In particular, the following query was submitted to the SPARQL endpoint of the Scottish portal to retrieve the compatible datasets.

```
PREFIX sdmx-dim:<http://purl.org/linked-
data/sdmx/2009/dimension#>
SELECT distinct ?b
WHERE {?a qb:dataSet  ?b;
       sdmx-dim:refPeriod
<http://reference.data.gov.uk/id/year/2017>.
       ?a sdmx-dim:refArea  [?m
<http://statistics.gov.scot/def/foi/collection/data-
zones-2011>].}
```

The query resulted in 17 datasets other than the “House Sales Prices” dataset. Each dataset includes one or more measures and thus the final list includes 23 variables that will be used in the creation of the predictive model. The final dataset containing the 23 variables was created by submitting a second query to the Scottish SPARQL endpoint. Table 1 presents the variables along with the results of an initial statistical analysis. Two Independent Sample t Test was used for the 23 continuous variables. Statistical analysis was performed using Python’s SciPy library and values of $p < 0.05$ were considered statistically significant.

Table 1. The data collected from the Scottish portal. Values are mean (± SD).

Variables	Overall (N = 5,841)	Above average (n = 2,340)	Below average (n = 3,501)	p value
Number of house sales	14.37 (11.36)	17.08 (15.39)	12.55 (6.99)	<0.001
Number of dwellings per hectare	19.91 (20.89)	15.99 (22.16)	22.53 (19.58)	<0.001
Number of flats	144.68 (168.62)	109.11 (162.31)	168.45 (168.59)	<0.001
Percentage of flats	33.43 (32.01)	25.02 (30.45)	39.05 (31.8)	<0.001
Median number of rooms in dwellings	3.99 (0.89)	4.52 (0.91)	3.63 (0.69)	<0.001
Number of dwellings	383.66 (117.88)	376.56 (126.75)	388.4 (111.32)	<0.001
Percentage of total dwellings that are long empty	1.37 (1.48)	1.41 (1.5)	1.34 (1.47)	0.08
Percentage of total dwellings that are occupied	96.12 (3.98)	95.81 (4.47)	96.33 (3.6	<0.001
Percentage of total dwellings that are second homes	0.93 (2.69)	1.43 (3.37)	0.59 (2.05)	<0.001
Percentage of population living proxime to derelict site	28.12 (39.05)	14.43 (28.62)	37.26 (42.28)	<0.001
Scottish access to bus indicator (weekday)	23.66 (36.73)	21.17 (40.97)	25.33 (33.5)	<0.001
Scottish access to bus indicator (weekend)	15.73 (25.22)	14.07 (27.6)	16.84 (23.44)	<0.001
Number of births	7.61 (5.3)	7.51 (6.56)	7.67 (4.25)	0.26
Number of deaths	8.51 (6.6)	7.56 (6.81)	9.15 (6.37)	<0.001
Mid-year population estimates	799.12 (221.37)	841.8 (267.66)	770.6 (178.49)	<0.001
Number of families receiving child benefit	80.15 (33.99)	72.77 (36.58)	85.08 (31.19)	<0.001
Number of children receiving child benefit	131.56 (58.55)	120.27 (61.9)	139.11 (54.93)	<0.001
Comparative illness factor	94.68 (50.64)	57.93 (28.25)	119.24 (47.32)	<0.001
Standardised mortality ratio	96.03 (42.38)	79.33 (39.34)	107.2 (40.64)	<0.001
Hospital stays related to alcohol misuse: standardized ratio	92.95 (88.95)	49.78 (48.9)	121.8 (97.61)	<0.001
Proportion of population being prescribed drugs for anxiety, depression or psychosis	18.56 (5.1)	15.2 (3.69)	20.81 (4.65)	<0.001
The number of people who are employment deprived	43.18 (33.98)	23.2 (19.37)	56.54 (35.07)	<0.001
The percentage of people who are employment deprived	8.58 (6.22)	4.35 (3.07)	11.41 (6.18)	<0.001

There are 2,340 data zones with mean house price above the average price. These areas are more likely to have dwellings with more rooms, dwellings that are not someone's main residence (second homes), healthier population, and to be more populated. In addition, they were less likely to have good accessibility to public transport, flats, families receiving child benefits, hospital stays related to alcohol misuse, people who are employment deprived, people being prescribed drugs for anxiety, and people living within 500 meters of a derelict site.

4.3 Create a Predictive Model

In this paper, XGBoost is used to create the predictive model. The dataset created in the previous step was split in train and test sets in order to ensure that the evaluation of the model is unbiased. We tune the model using the train set and then we evaluate the final model in the test set. A total of 4,380 (75%) and 1,461 (25%) data zones were randomly assigned to the train and test sets, respectively. Repeated k-fold Cross-Validation was employed in order to ensure that our model will have low variance and bias. Part of data (the training sample) is used for training each algorithm, and the remaining part (the validation sample) is used for estimating the risk of the algorithm. Using Cross-Validation, several XGBoost parameters were selected to maximize model performance.

The performance of the models was assessed by measuring the total area under the receiver-operating curve (AUC). The model that was created and tuned is applied to the test set to get the AUC score in the holdout data.

The results of the machine learning model creation along with the optimal hyper-parameter values selected are depicted in Table 2. The holdout AUC score is 0.91.

Table 2. The AUC score to the created model along with the optimal hyper-parameter values

Train set AUC	Test set AUC	XGBoost parameter values
0.915	0.91	'colsample_bylevel': 0.8, 'colsample_bytree': 0.6, 'learning_rate': 0.1, 'max_depth': 3, 'n_estimators': 400, 'subsample': 0.8

4.4 Explain the Predictive Model

In this study the following types of visualizations are employed to explain the created predictive model:

- SHAP summary plots: beeswarm plots where the dot's position on the y-axis is determined by the feature and on the x-axis by the Shapley value. The color represents the value of the feature from low to high.
- SHAP Dependence Plots show how a feature's value (x-axis) impacts the prediction (y-axis) of every sample (each dot) in a dataset

- SHAP Interaction Value Dependence Plots: dependence plot on the SHAP interaction values, which allows to separately observe the main effects and the interaction effects.

SHAP Summary Plots. In Fig.1 the SHAP summary plot is presented in the form of a set of beeswarm plots. The order of the features reflects their importance, i.e., the sum of the SHAP value magnitudes across all samples. Each point on the summary plot is a Shapley value for a feature and an instance. The position on the y-axis is determined by the feature and on the x-axis by the Shapley value. The color represents the value of the feature from low to high.

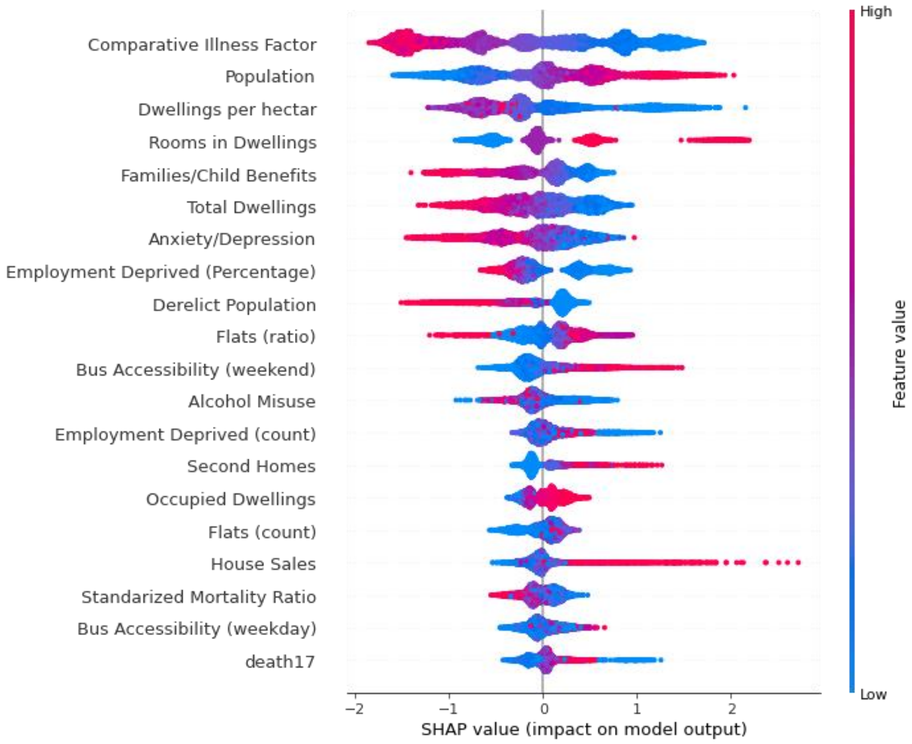

Fig. 1. A set of beeswarm plots, where each dot corresponds to an individual data zone in the study.

The plot reveals that the Comparative Illness Factor (CIF) is the most important feature globally. CIF is a measure of chronic health conditions that takes account of people from all ages. CIF greater than 100 indicates poorer health conditions relative to Scotland and vice-versa. The plot indicates the direction of the effects, meaning, for example, that low CIF data zones (blue) have higher probablity of having expensive houses than high CIF data zones (red). Moreover, the plot presents the distribution of effect sizes, such as the long tails of many variables. These long tails mean that features with a low global importance can be extremely important for specific data zones. For example, although the number of house sales normally do not imply the level of house prices, in some abnormal cases the high number of sales indicate areas with expensive houses.

SHAP Dependence Plot. The impact of a feature's value to the prediction can be revealed by using the SHAP dependence plots. Fig. 2-a clearly reveals the inflection point on the impact of the comparative illness factor (CIF) to the house prices. For low CIF values overall SHAP values are positive up to a point around 75. Then, SHAP values are negative, which means that by increasing CIF, the probability of high house prices decreases.

The vertical dispersion of SHAP values at a single feature value is driven by interaction effects. In Fig. 2-b, the number of people who are employment deprived is chosen for coloring to highlight possible interactions. The blue points mostly appear for lower values of CIF, meaning that areas of poorer health conditions tend to have more unemployed people.

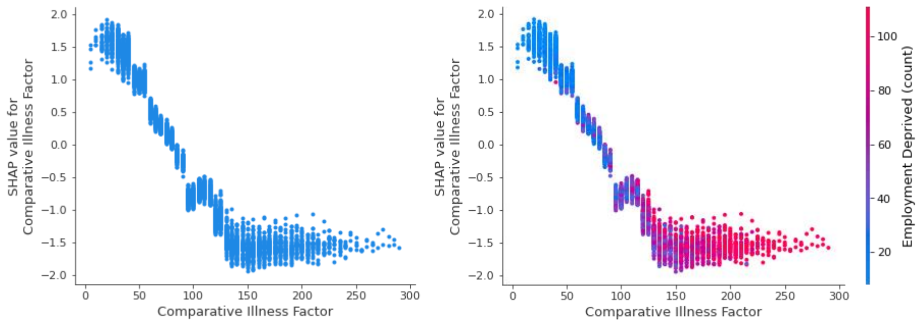

Fig. 2. (left) SHAP dependence plot of comparative illness factor vs. its SHAP value in the created predictive model, (right) SHAP dependence plot of comparative illness factor with interaction visualization with the number of people who are employment deprived.

SHAP Interaction Value Dependence Plot. SHAP interaction values can be interpreted as the difference between the SHAP values for feature *i* when feature *j* is present and the SHAP values for feature *i* when feature *j* is absent [18]. The interaction effect is the additional combined feature effect after accounting for the individual feature effects. In this sub-section SHAP interaction effects are explored. Towards this end, the plots of the SHAP interaction values of multiple pairs of variables are created and presented in Fig. 3.

The plot of the SHAP interaction value of 'Total Dwellings' with 'Population' (Fig.3-a) shows how the effect of total number of dwellings on the probability of high house prices varies with population. The plot of the SHAP interaction value of 'Anxiety/Depression' (i.e., proportion of population being prescribed drugs for anxiety, depression or psychosis) with 'Population' (Fig.3-b) shows that in data zones with depressed population of more than 23%, total population has a different effect on the probability of high house prices depending on the size of the total population. Small population size has negative effect, while large population size has positive effect. Moreover, the plot of SHAP interaction value of the percentage of flats with the comparative illness factor (CIF) (Fig.3-c) shows that in data zones of poor health condition (CIF > 100) the effect of the percentage of flats reverses at a point around 25%. Similar patterns can be found in three more plots. Fig.3-d shows that the effect of 'dwellings per hectar' is different in data zones with small and large number of births. The same holds in Fig. 3-e and Fig.3-f that depict the effect of house sales in data zones with high and low number of dwellings and dwellings per hectar respectively.

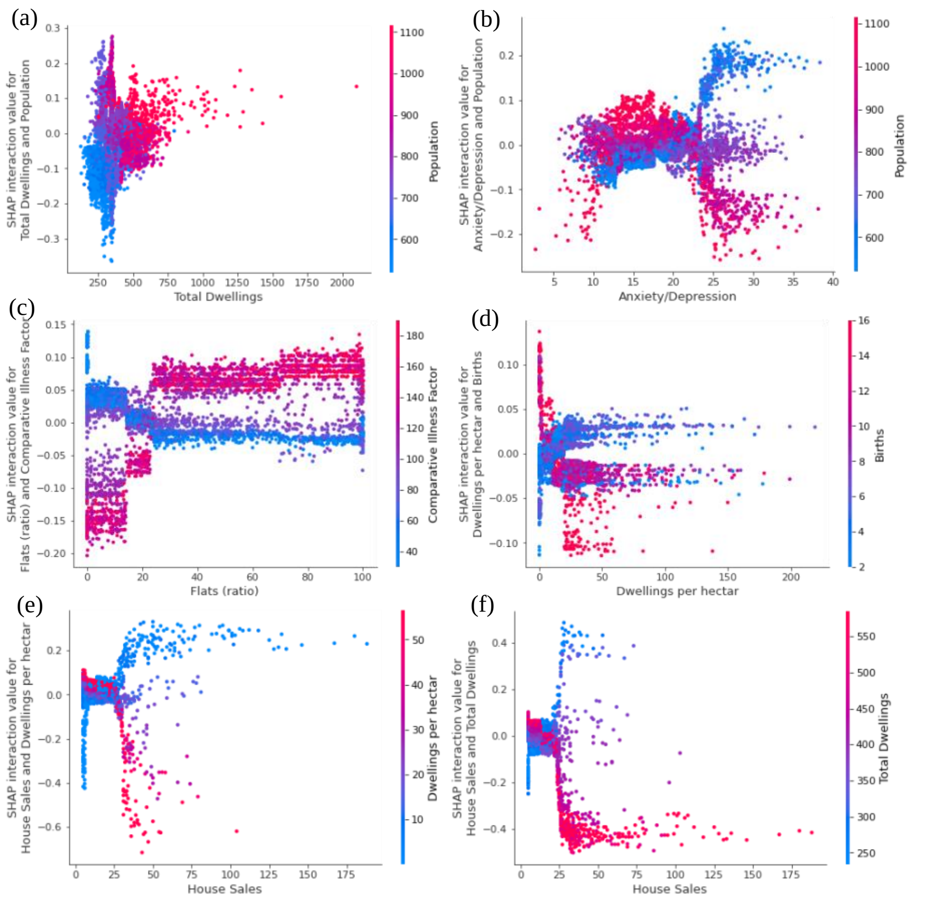

Fig. 3. SHAP interaction value dependence plots

5 Conclusion

Machine learning and artificial intelligence models promise to streamline public services and policy making. In many cases these models have been proven accurate to predict the outcome of relevant events and thus effective to support public administration and policy makers. However, there is growing emphasis on building tools and techniques for explaining these models in an interpretable manner. The objective of this paper is to illustrate the applicability and the value of applying eXplainable Artificial Intelligence (XAI) techniques on open government data that are formulated as linked open statistical data. Towards this end, a case study using data from the official open data portal of the Scottish Government is presented. In this case, an XGBoost algorithm is used to create and SHAP framework to explain a model that predicts the probability the mean house price of a data zone to be above the average price across all 6,976 data zones. The AUC score of the created model was 0.91, while the analysis based on Shapley values revealed some interesting insights.

The analysis demonstrated that linked data facilitate the discovery and collection of high-quality data. The definition and submission of two SPARQL queries was sufficient to create the final dataset that was used for the creation of the predictive model. The first query identified 17 compatible and thus candidate datasets, while the second revealed 23 variables populated with data for the 2011 data zones. An important direction for future research in this area is to combine datasets that span across different data portals and countries. Interoperability challenges across data portals need to be further analyzed and addressed.

The creation of an accurate model using an advanced tree-based ensemble algorithm demonstrates that open government data can be used in machine learning scenarios. In this direction, the specification of an appropriate question to answer is of vital importance.

The local explanation analysis using Shapley values demonstrated the importance of such an analysis in policy making and/or public administration context. The computation of a feature's effect per individual case (e.g., in each data zone area in our case study) enables applying different policies or making different decisions based on the distinct characteristics of each case. For example, although in general the high proportion of population being prescribed drugs for anxiety, depression or psychosis indicates areas with low house prices, in few areas this high proportion abnormally contributes to predict an area of high prices. This significantly improves the ability of public administrations and policy makers to make more accurate data-driven decisions and apply more focused evidence-based policies.

Acknowledgements. This publication has been produced in the context of the EU H2020 Project inGOV which is co-funded by the European Commission under the Grant agreement ID: 962563.

References

1. Arrieta, A. B., Díaz-Rodríguez, N., Del Ser, J., Bennetot, A., Tabik, S., Barbado, A., ... Herrera, F.: Explainable Artificial Intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI. Information Fusion, 58, 82-115 (2020)
2. Chen, T., Guestrin, C.: XGBoost: A Scalable Tree Boosting System. In Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD '16). Association for Computing Machinery, New York, NY, USA, 785–794 (2016)

3. Deeks, A.: The Judicial Demand For Explainable Artificial Intelligence. Columbia Law Review, 119(7), 1829-1850 (2019)
4. Freund, Y., Schapire, R. E.: A decision-theoretic generalization of on-line learning and an application to boosting. Journal of computer and system sciences 55(1), 119-139 (1997)
5. Friedman, J.: Greedy Function Approximation: A Gradient Boosting Machine. *The Annals of Statistics*, *29*(5), 1189-1232 (2001)
6. Gao Y., Janssen M.: Generating Value from Government Data Using AI: An Exploratory Study. In: Viale Pereira G. et al. (eds) Electronic Government. EGOV 2020. Lecture Notes in Computer Science, vol 12219. Springer, Cham (2020)
7. Gunning, D.: Explainable artificial intelligence (xai). Defense Advanced Research Projects Agency (DARPA), nd Web, 2(2) (2017)
8. Janssen M, Hartog M, Matheus R, Yi Ding A, Kuk G.: Will Algorithms Blind People? The Effect of Explainable AI and Decision-Makers' Experience on AI-supported Decision-Making in Government. Social Science Computer Review (2021). doi:10.1177/0894439320980118
9. Kalampokis, E., Karamanou, A., Tarabanis, K.: Interoperability Conflicts in Linked Open Statistical Data. Information, 10(8), 249 (2019)
10. Kalampokis, E., Tambouris, E., Tarabanis, K.: Linked open cube analytics systems: potential and challenges. IEEE Intell. Syst. 31(5), 89–92 (2016)
11. Kalampokis, E., Zeginis, D., Tarabanis, K.: On modeling linked open statistical data. Journal of Web Semantics 55, 56-68 (2019)
12. Kouziokas, G.N.: The application of artificial intelligence in public administration for forecasting high crime risk transportation areas in urban environment. Transp. Res. Procedia 24, 467–473 (2017)
13. Lundberg S.M. et al.: Explainable machine learning predictions to help anesthesiologists prevent hypoxemia during surgery. Nature Biomedical Engineering 2, 749–760 (2018)
14. Lundberg, S. M., Erion, G., Chen, H., DeGrave, A., Prutkin, J. M., Nair, B., ... Lee, S. I.: From local explanations to global understanding with explainable AI for trees. Nature machine intelligence, 2(1), 2522-5839 (2020)
15. Lundberg, S.M.: Erion, G.G. and Lee, S.I.: Consistent individualized feature attribution for tree ensembles. *arXiv preprint arXiv:1802.03888* (2018)
16. Lundberg, S., Lee, S.I.: A unified approach to interpreting model predictions. (2017) arXiv preprint arXiv:1705.07874
17. Montavon, G., Samek, W., Müller, K. R.: Methods for interpreting and understanding deep neural networks. Digital Signal Processing, 73, 1-15 (2018)
18. Parsa, A. B., Movahedi, A., Taghipour, H., Derrible, S., Mohammadian, A. K.: Toward safer highways, application of XGBoost and SHAP for real-time accident detection and feature analysis. Accident Analysis & Prevention, 136, 105405 (2020)
19. Ramon, Y., Martens, D., Provost, F. et al.: A comparison of instance-level counterfactual explanation algorithms for behavioral and textual data: SEDC, LIME-C and SHAP-C. Adv Data Anal Classif 14, 801–819 (2020)
20. Ribeiro MT, 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 (2016)
21. Roscher, R., Bohn, B., Duarte, M. F., Garcke, J.: Explainable machine learning for scientific insights and discoveries. IEEE Access, 8, 42200-42216 (2020)
22. Saltos, G., Cocea, M.: An exploration of crime prediction using data mining on open data. Int. J. Inf. Technol. Dec. Making 16(5), 1155–1181 (2017)
23. Shapley, L.S.: A value for n-person games. Contrib. to Theory Games, 307–317 (1953).

24. Slack, D., Hilgard, S., Jia, E., Singh, S., Lakkaraju, H.: Fooling lime and shap: Adversarial attacks on post hoc explanation methods. In Proceedings of the AAAI/ACM Conference on AI, Ethics, and Society, pp. 180-186 (2020)
25. de Sousa, W. G., de Melo, E. R. P., Bermejo, P. H. D. S., Farias, R. A. S., Gomes, A. O.: How and where is artificial intelligence in the public sector going? A literature review and research agenda. Government Information Quarterly, 36(4), (2019)