# Predicting Digital Winners and Losers in Economic Crises Using Artificial Intelligence and Open Government Data

Euripidis Loukis [1[0000-0002-5932-4128]] and Mohsan Ali[1[0000-0002-3956-3543]]

[1] University of Aegean, Samos, Greece
[eloukis, Mohsan]@ aegean.gr

**Abstract.** In market-based economies often appear significant decreases of economic activity, which lead to recessionary economic crises. These economic crises have quite negative consequences for firms, as they lead to significant decrease of their sales revenues; firms respond by decreasing on one hand their production and in general operational activities and expenses, personnel employment and materials' procurement, and on the other hand their investments in production equipment, digital technologies, etc., which leads to technological obsolescence. This reduction of investments, and especially of the ones in digital technologies, due to their importance for firms' efficiency, effectiveness, and innovation, can have quite negative impact on their future competitiveness, and even put at risk their survival. However, these negative consequences of economic crises differ significantly among firms: some of them exhibit a lower vulnerability to the crisis, so they have less negative consequences, while some other firms exhibit a higher vulnerability, and have more negative consequences; so the competitive position of the former is significantly strengthened with respect to the latter, and finally the former are the 'winners' of the crisis, while the latter are the 'losers'. This paper proposes a methodology for predicting the winner and loser firms of future economic crises with respect to a highly important class of technologies: the digital technologies. In particular, the proposed methodology enables the prediction of the multi-dimensional 'pattern of digital vulnerability' of an individual firm to a future economic crisis, which consists of the degrees of reduction of the main types of 'digital investments' as well as 'digital operating expenses' in a future economic crisis. For this purpose, we are using Machine Learning algorithms, in combination with the Synthetic Minority Oversampling Technique (SMOTE), in order to increase their performance, which are trained using open government data from Statistical Authorities. Furthermore, a first application/validation of the proposed methodology is presented, using open data from the Greek Statistical Authority for 363 firms for the severe Greek economic crisis period 2009-2014, which gave satisfactory results concerning the prediction of nine different aspects of digital vulnerability to economic crisis (five of them concerned the main types of digital investment, and the other four concerned the main types of digital operating expenses).

**Keywords:** Economic Crises, Digital Technologies, Artificial Intelligence, Machine Learning

## 1 Introduction

In market-based economies often appear significant decreases of economic activity, which lead to recessionary economic crises of different intensities, geographical scopes and durations [1-6]. In [3] are briefly described numerous economic crises that appeared during the previous century, as well as their origins and negative consequences for the society and the economy; in the beginning of this century initially we had the severe 2007 Global Financial Crisis [1, 7], shortly after the end of it we had the COVID-19 that gave rise to another economic crisis [8], while recently the Ukraine war resulted in big increases in the prices of oil, gas, wheat and other goods, which spark another economic crisis.

These economic crises have quite negative consequences for firms, as they lead to decrease of the demand for their products and services, and therefore of their sales; firms respond to this by decreasing on one hand their production and in general operational activities and expenses, as well as personnel employment and materials' procurement, and on the other hand their investment in production equipment, digital technologies, etc., which leads to technological obsolescence. This reduction of investment, and especially of the ones in digital technologies, due to their high importance for firms' efficiency, effectiveness, and innovation, can have quite negative impact on the future competitiveness of firms, or even put at risk the survival of some of them [1-2, 9-12]. However, these negative consequences of economic crises differ significantly among firms: some of them exhibit a higher capacity to cope with the crisis, so they have less negative consequences, and therefore higher resilience to the crisis; on the contrary, some other firms exhibit a lower capacity to cope with the crisis, so they have more negative consequences, and therefore lower resilience to the crisis [1-3, 13]. This results in a strengthening of the competitive position of the former with respect to the latter, so finally the former are the 'winners' of the crisis, while the latter are the 'losers'. Governments, in order to mitigate the negative consequences of these economic crises for the firms and the citizens undertake huge interventions, such as large-scale economic stimulus programs, which include the provision to firms of tax rebates, financial assistance, subsidies, financial support for investments, low-interest (or even no-interest) loans, etc. [14-17].

This paper proposes a methodology for predicting the winner and loser firms of future economic crises with respect to a highly important class of technologies for firms: the digital technologies. In particular, the proposed methodology enables predicting the multi-dimensional 'patterns' of digital vulnerability to economic crisis of individual firms. This digital vulnerability pattern of a firm includes the degrees of reduction of the main types of 'digital investment', e.g., in ICT hardware, software, staff training, etc., due to economic crisis. Furthermore, since the benefits that a firm obtains from the use of digital technologies depends not only on the magnitude of the above digital investments, but also on its relevant 'digital operational expenses', e.g., for ICT personnel, for cloud computing, etc. So, the pattern of digital vulnerability to economic crisis of an individual firm has to include also the degrees of reduction of the main types of 'digital operating expenses'.

For this purpose, we employ Artificial Intelligence (AI) algorithms from the area of Machine Learning (ML) [18-20], which are used in order to construct a set of prediction models of the digital vulnerability of a firm concerning the abovementioned main types of digital investment as well as digital operating expenses; as independent variables are used the characteristics of each individual firm (e.g. concerning human resources' size and quality (education and skills), production equipment, use of ICT, innovation, exports, strategic directions, comparative financial performance vis-à-vis competitors, etc.). For the training of these prediction models are used open government data (OGD) [21-23] from Statistical Authorities. The great potential of using AI in order to make advanced analysis of OGD, and extract from them valuable knowledge and prediction models, which can be quite useful for addressing the serious challenges that modern societies face, is highlighted and analyzed in a recent report of UNESCO [24]; in this report are also suggested some guidelines for governments in order to promote and accelerate this advanced exploitation of OGD through AI. However, at the same this report mentions that the use of OGD for training AI models can pose important challenges, which concern data privacy, data ethics, legal limitations, data infrastructures, data governance, as well as unrepresentative datasets with limited data from marginalized groups.

Furthermore, in this paper a first application/validation of the proposed methodology is presented, using OGD from the Greek Statistical Authority for 363 firms for the severe Greek economic crisis period 2009-2014, which gave satisfactory results concerning the prediction of nine different aspects of digital vulnerability to economic crisis (five of them concerned the main types of digital investment, and the other four concerned the main types of digital operating expenses).

The proposed methodology can be useful for government agencies that design and implement various types of interventions, such as large-scale economic stimulus programs, for mitigating the negative consequences of economic crises, in order to focus their digital actions (e.g., financial support or subsidies for firms' digital investment, employment of ICT personnel or digital operating expenses in general) on the most digitally vulnerable firms, which most need government support. Also, our methodology can be useful for banks and investment firms, in order to grant loans to and invest in firms that will not show high digital vulnerability to economic crisis, and therefore are not at risk of digital obsolescence (which can have negative consequences for their future competitiveness or even for their survival), in a future economic crisis. Finally, our methodology can be useful for ICT firms as well as consulting firms offering ICT-related services during economic crises, enabling them to identify firms that will not show high digital vulnerability to the crisis, in order to focus on them their marketing and sales activities (thus increasing the likelihood of substantial sales revenue).

The remaining of the paper has been structured as follows: the second section explains the background of our research; the third section describes the proposed methodology, and then the fourth section presents the abovementioned application of it; lastly, the conclusions are summarized in the fifth section.

## 2 Background – AI in Government

Though AI remained for long time mainly in a research stage in university research labs, in the last decade there has been a sharp increase in its 'real-life' application and exploitation, initially by private sector firms, and recently by government agencies as well [25-30]. The huge amounts of data possessed by government agencies motivated them to proceed to a more advanced and sophisticated exploitation of them using AI techniques, mainly from the area of ML, in order to automate, support or enhance/augment more sophisticated mental tasks than the simpler routine ones that are currently automated, supported or enhanced/augmented by the traditional information systems of government agencies.

There have been some first cases of successful AI use in various domains of government activity, for instance in tax administration for the fight against tax evasion (e.g. for identifying citizens and firms with high tax evasion risk, in order to conduct more targeted tax audits); in healthcare for supporting diseases' diagnosis and prevention, as well as treatment planning and support of clinical decision making; in social policy for fraud detection, for the prediction of higher-risk youth with respect to future criminal activity, in order to target prevention interventions; in policing for the prediction of persons or geographical areas with high criminal activity risk (e.g. forecast spatial crime patterns based on socio-demographic factors for directed patrolling in order to increase effectiveness in crime prevention), or for face recognition from images or video data (e.g., from cameras); for improving the interaction with citizens through chatbots, etc. With respect to the main directions/objectives of the first AI uses in government a recent relevant study of the AI Watch of the European Union [28] concludes that most cases of AI use in the public sector of the member states aim at the improvement of public services (e.g., through the provision of personalized and citizen-centered information and services), followed by the improvement of internal administrative efficiency (e.g., through improved management of public resources as well as increase of the quality and decrease of the cost of internal processes). Furthermore, there is an intense debate concerning the role that AI can play in government; we can distinguish two main roles for AI [30], which correspond to two quite different approaches to/ways of exploiting AI in government:

a) Use of AI for automating some administrative tasks (in order to achieve cost savings and efficiencies, which is quite useful due to the fiscal constraints that most governments face but might lead to public sector workforce substitution and job losses.

b) AI use for augmenting some administrative tasks, mainly higher-level ones associated with decision-making and policy design (by providing to them some additional useful inputs, which can increase the quality and the effectiveness of the execution of these tasks).

Comprehensive reviews of the research that has been conducted concerning the use of AI in government are provided in [25], [27], [29] and [30].

However, it is widely recognized that the research and practice in the area of AI use in government has exploited only a small part of the large potential that the wide spectrum of AI techniques have for automating, supporting or enhancing/augmenting various government functions and activities; also, most cases of AI use in government concern low or medium importance and financial magnitude activities [25], [27] and [30]. Therefore, extensive further research is required in order to exploit more this large potential, developing new ways and methodologies of using AI in government, especially in its more important and costly activities for increasing their efficiency and effectiveness. Also, further research is required also concerning the use of AI for analyzing OGD in order to achieve higher levels of exploitation of them, by extracting from them valuable knowledge and prediction models, which can be useful for both public as well as private sector decision-making, following the recent recommendations of the UNESCO [24]. The present study makes a contribution in both these directions, by developing a methodology of using AI/ML for making an advanced exploitation of OGD, for the development of prediction models in order to support and enhance/augment one of the most important and costly kinds of interventions that governments have to make: the interventions and especially the large-scale economic stimulus programs they undertake in tough times of economic crises in order to reduce their negative consequences, focusing on the 'digital' interventions.

## 3 Proposed Methodology

The proposed methodology aims to predict the multi-dimensional 'pattern of digital vulnerability' to an economic crisis (DVEC) of an individual firm, which is defined as a vector that has as components the degrees of reduction of the main types of digital investments (e.g. in hardware, software, etc.) as well as digital operating expenses (e.g. for ICT staff payroll, external ICT services); they can be measured either in a continuous scale or in a 5-levels Likert scale: not at all, small, moderate, large, very large), during an economic crisis (Figure 1):

$$DVEC = [DVEC_1, DVEC_{2, \ldots\ldots,} DVEC_N]$$

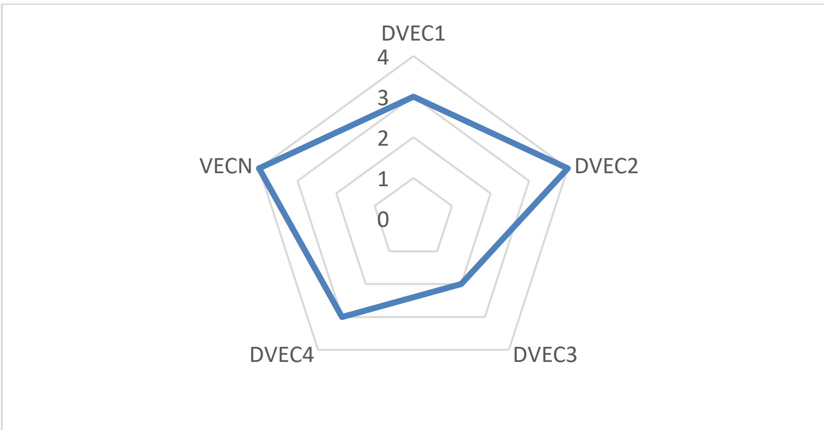


**Fig. 1**. Multidimensional Pattern of Firm's Digital Vulnerability to Economic Crisis

Firms predicted to have a high degree of reduction of the above main types of digital investments as well as digital operating expenses in a future economic crisis have a risk of becoming ‘digitally obsolete’; they will be the ‘digital losers’ of the crisis, and this might have negative impact on their future competitiveness, or even put at risk their survival. On the contrary, firms predicted to have a low degree of reduction of the above main types of digital investments and digital operating expenses in a future economic crisis, will be the ‘digital winners’ of the crisis, and this will strengthen their competitive position with respect to the ‘digital losers’.

For each of the abovementioned components/dimensions of firm’s digital vulnerability to economic crisis $DVEC_1$, $DVEC_{2, \ldots\ldots,}$ $DVEC_N$ a prediction model of it is constructed, having it as dependent variable. In order to determine the appropriate independent variables of these prediction models we have been based on management science research, which has been developed several frameworks of the main elements of a firm that determine its performance; the ‘Leavitt’s Diamond’ framework is the most widely recognized one, which includes five main elements: strategy, processes, people, technology, and structure [31-32]. We can expect that these five main elements of the ‘Leavitt’s Diamond’ framework will be the main determinants of the performance of a firm not only in normal economic periods but also in economic crisis ones as well.

So, the prediction models of firm’s digital vulnerability concerning the main types of digital investments and digital operating expenses $DVEC_i$ will include five corresponding groups of independent variables concerning:

a) strategy (e.g., degree of adoption of the main competitive advantage strategies defined in relevant strategic management literature [33], such as cost leadership, differentiation, focus, innovation, etc.)

b) processes (e.g., main characteristics of firm’s processes, such as complexity, flexibility, etc.)

c) people (e.g., shares of firm’s human resources having different levels of education or specific skills, certifications, etc.)

d) technology (e.g., use of various production technologies, digital technologies, etc.)

e) structure (e.g., main characteristics of firm’s structure, degree of adoption of ‘organic’ forms of work organization, such as teamwork, etc.)

and also, a sixth group of independent variables concerning general information about the firm, such as size, sector, comparative performance vis-à-vis competitors, etc.

For the construction of each of these prediction models we can use/try the main supervised ML algorithms described in relevant literature [18-20], such as Decision Tree (DT), Random Forest (RF), Logistic Regression (LR), Support Vector Machines (SVM), and Multilayer Perceptron (MLP), compare their prediction performances, and finally select the one with the highest prediction performance. For training them we can use relevant OGD for economic crisis periods provided by Statistical Authorities; the available dataset is divided into two parts: the ‘training dataset’, which is used for constructing the prediction model, and the ‘test dataset’, in which the prediction performance of this model is evaluated, by calculating its prediction accuracy, precision, recall, and F-Score are calculated. The formals for these terms are given below in the equations 1, 2, 3, 4; they are calculated based on the numbers of TP (True positive), TN (True negative), FP (False positive), and FN (False negative).

$$Accuracy = \frac{TN+TP}{TN+FP+TP+FN} \tag{1}$$

$$Precision = \frac{TP}{TP+FP} \tag{2}$$

$$Recall = \frac{TP}{TP+FN} \tag{3}$$

$$F - score = \frac{TP}{TP+FN} \tag{4}$$

However, the abovementioned datasets we use for the construction of the models usually a) are too noisy because of the missing values, b) their size can be not large enough to train a healthy and accurate ML model, and c) are unbalanced with respect to the classes (they include small numbers of observations/samples for some of the classes, and much larger numbers of observations/samples for some other classes); these can result in prediction models with lower prediction performance and also biased. In order to address problems b) and c) our methodology includes a pre-processing of these datasets using the Synthetic Minority Oversampling Technique (SMOTE) [34]; this technique increases the number of samples of the dataset using the existing samples of the classes (oversampling), balances the dataset with respect to the number of samples of each class, fills missing values, which enable the estimation of better prediction models with higher prediction performance. Furthermore, our methodology includes an initial Exploratory Data Analysis (EDA) in order to get a first insight of the data through visualizations, and make some necessary transformations, and then a Principal Component Analysis (PCA) in order to analyze the importance of the features, and select the eliminate the important ones, and eliminate the ones that are not important, which helps to improve performance and reduce training time.

# 4 Application

A first application/validation of the proposed methodology was made using a dataset that was released by the Greek Statistical Authority on request by the authors, and after signing an agreement concerning its use, so it constitutes OGD freely available (under specific terms) for research purposes. The full dataset was comprised of 363 instances, with each instance being an independent firm. It included for each firm the following features/variables:

a) Nine variables concerning different dimensions of firm's digital vulnerability to the severe economic crisis that Greece experienced between 2009 and 2014 (dependent variables). Five of them concerned digital investment: degree of reduction of the total hard and soft ICT investment (in ICT hardware, ICT software, ICT training of staff and ICT consulting), ICT hardware, ICT hardware, ICT staff training and ICT consulting; and the remaining four concerned digital operating expenses: degree of reduction of total ICT operating expenses (for ICT personnel and ICT services), ICT staff expenses, ICT outsourcing expenses and cloud computing services' expenses. All these variables were measured in a 6-levels Likert scale (increase, negligible impact, small decrease, moderate decrease, large decrease, very large decrease), and were then converted to binary ones (with the first three values being converted to 'non-vulnerable' and the

other three being converted to 'vulnerable'). The corresponding questions of the Greek Statistical Authority questionnaire are shown in the Appendix.
b) Forty variables concerning various firm's characteristics with respect to strategy, personnel, technology (focusing on the use of various digital technologies), structure (focusing on the use of organic' forms of work organization, such as teamwork) and also some general information about the firm (size measured through the number of employees, sector (services or manufacturing), comparative financial performance in the last three years in comparison with competitors) (independent variables).

The dataset was noisy because of missing values and was converted to noise free by filling the missing value with the most suitable value based on the type of each variable (for instance, if the variable was ordinal, then the missing values were filled with the highest relative frequency value of this variable). Then exploratory data analysis (EDA) was applied in order to gain a better insight into the data through visualizations and make some necessary transformations. As a next step, since the size of our data set (which included data for 363 firms as mentioned above) did not allow constructing/training a good, supervised ML prediction model (with good prediction performance), we used the abovementioned (in section 3) oversampling and class-balancing algorithm SMOTE. Finally, the dataset was divided into a training dataset including 66% of the samples and a testing dataset including 33% of the samples. The former was used for the training of ML models with the following algorithms: Decision Tree (DT), Random Forest (RF), Logistic Regression (LR), Support Vector Machines (SVM), and Multilayer Perceptron (MLP). The latter was used for assessing the prediction performance of these ML models; we can see the results (prediction accuracy, precision, recall and F-score in Table 1 (with bold are shown for each dimension of digital crisis vulnerability the results for the best performing algorithm).

Table 1: Prediction Performance of AI/ML Algorithms for each Digital Crisis Vulnerability Dimension

| Dimension of Digital Crisis Vulnerability | AI /ML Algorithm | Accuracy | Precision | Recall | F1-score |
|---|---|---|---|---|---|
| Degree of Reduction of Total Hard and Soft ICT Investment | Decision Tree | 0.76 | 0.76 | 0.76 | 0.76 |
| | Random Forest | 0.84 | 0.84 | 0.84 | 0.83 |
| | Logistic Regression | 0.67 | 0.65 | 0.67 | 0.65 |
| | **SVM** | **0.86** | **0.86** | **0.86** | **0.86** |
| | MLP | 0.71 | 0.70 | 0.71 | 0.69 |
| Degree of Reduction of Investment in ICT Hardware | Decision Tree | 0.78 | 0.79 | 0.78 | 0.78 |
| | **Random Forest** | **0.88** | **0.89** | **0.88** | **0.87** |
| | Logistic Regression | 0.76 | 0.75 | 0.76 | 0.74 |

| | | | | | |
|---|---|---|---|---|---|
| | SVM | 0.86 | 0.86 | 0.86 | 0.85 |
| | MLP | 0.72 | 0.70 | 0.72 | 0.68 |
| Degree of Reduction of Investment in ICT Software | Decision Tree | 0.77 | 0.77 | 0.77 | 0.77 |
| | **Random Forest** | **0.87** | **0.88** | **0.87** | **0.86** |
| | Logistic Regression | 0.67 | 0.64 | 0.67 | 0.65 |
| | SVM | 0.87 | 0.88 | 0.87 | 0.86 |
| | MLP | 0.71 | 0.77 | 0.71 | 0.62 |
| Degree of Reduction of Investment ICT Training of Staff | Decision Tree | 0.80 | 0.80 | 0.80 | 0.80 |
| | **Random Forest** | **0.90** | **0.91** | **0.90** | **0.90** |
| | Logistic Regression | 0.72 | 0.70 | 0.72 | 0.70 |
| | SVM | 0.90 | 0.90 | 0.90 | 0.90 |
| | MLP | 0.76 | 0.75 | 0.76 | 0.75 |
| Degree of Reduction of Investment ICT Outsourcing | Decision Tree | 0.78 | 0.78 | 0.78 | 0.78 |
| | **Random Forest** | **0.85** | **0.85** | **0.85** | **0.85** |
| | Logistic Regression | 0.74 | 0.73 | 0.74 | 0.73 |
| | SVM | 0.85 | 0.85 | 0.85 | 0.84 |
| | MLP | 0.74 | 0.73 | 0.74 | 0.73 |
| Degree of Reduction of Total ICT Operating Expenses | Decision Tree | 0.76 | 0.76 | 0.76 | 0.76 |
| | **Random Forest** | **0.90** | **0.90** | **0.90** | **0.89** |
| | Logistic Regression | 0.73 | 0.72 | 0.73 | 0.72 |
| | SVM | 0.85 | 0.85 | 0.85 | 0.84 |
| | MLP | 0.72 | 0.71 | 0.72 | 0.68 |
| Degree of Reduction of ICT Staff Expenses | Decision Tree | 0.79 | 0.79 | 0.79 | 0.79 |
| | Random Forest | 0.91 | 0.91 | 0.91 | 0.91 |
| | Logistic Regression | 0.74 | 0.73 | 0.74 | 0.73 |
| | **SVM** | **0.91** | **0.91** | **0.91** | **0.91** |

| | | | | | |
|---|---|---|---|---|---|
| | MLP | 0.77 | 0.76 | 0.77 | 0.76 |
| Degree of Reduction of ICT Outsourcing Expenses | Decision Tree | 0.76 | 0.77 | 0.76 | 0.76 |
| | Random Forest | 0.88 | 0.88 | 0.88 | 0.87 |
| | Logistic Regression | 0.75 | 0.74 | 0.75 | 0.74 |
| | **SVM** | **0.88** | **0.88** | **0.88** | **0.88** |
| | MLP | 0.72 | 0.70 | 0.72 | 0.70 |
| Degree of Reduction of Cloud Computing Services Expenses | Decision Tree | 0.81 | 0.81 | 0.81 | 0.81 |
| | **Random Forest** | **0.92** | **0.92** | **0.92** | **0.92** |
| | Logistic Regression | 0.70 | 0.69 | 0.70 | 0.69 |
| | SVM | 0.89 | 0.90 | 0.89 | 0.89 |
| | MLP | 0.78 | 0.77 | 0.78 | 0.77 |

We can see that for all dimensions of firm's digital vulnerability to economic crisis we have high prediction performances, with prediction accuracies (for the best performing AI/ML algorithm for each dimension) ranging between 0,85 and 0,92. Overall, the results of this first application of the proposed methodology (concerning prediction performances) can be regarded as quite satisfactory, taking into account the small size of the dataset we have used (data from 363 firms), and provide a first validation of this methodology. We expect that using a larger dataset (as governments have such data for quite large numbers of firms) will allow training crisis-vulnerability prediction models with higher prediction performances.

## 5 Conclusion

In the previous sections we have described a methodology for predicting the whole pattern of digital vulnerability to economic crisis of individual firms with respect to the main types of digital investments as well as digital operating expenses. It is based on AI/ML techniques, in combination with SMOTE in order to increase their performance, which are trained using OGD from Statistical Authorities. This is in line with the recent recommendations of UNESCO for advanced exploitation of OGD using AI [24]. Our methodology enables a prediction of the degree of reduction of the main types of digital investment as well as digital operating expenses of an individual firm in a future economic crisis, based on its characteristics with respect to human resources, technologies, strategic orientations, processes and structure; this enables predicting the digital winner as well as the digital loser firms of future economic crises. The proposed methodology has as theoretical foundation the widely recognized ‘Leavitt’s Diamond’ framework [31-32] from management science.

Furthermore, a first application/validation of the methodology has been presented, using OGD from the Greek Statistical Authority about 363 firms for the severe Greek economic crisis period 2009-2014, which gave quite satisfactory results concerning the prediction of nine different aspects of digital vulnerability to economic crisis: five of them concerned the main types of digital investment, and the other four concerned the main types of digital operating expenses.

The research presented in this paper has some interesting implications for research and practice. With respect to research, it makes a significant contribution to two highly important research streams. First, it makes a contribution to the growing research stream concerning the use of AI in government, by developing a novel approach for a highly beneficial use of AI/ML for supporting and enhancing/augmenting a critical activity of government, which is characterized by quite high economic/social importance and financial magnitude: the interventions, and especially the large-scale economic stimulus programs, for mitigating the negative consequences of economic crises (focusing on the digital interventions/actions of these programs). Second, it makes a contribution to the OGD research stream, by providing an approach for increasing the economic/social value generation form the OGD through advanced processing of them using AI/ML techniques. With respect to practice, it provides a useful tool for government agencies responsible for the above interventions and programs that aim to mitigate the negative consequences of economic crises, as well as for banks and investment firms, for enhancing/augmenting their process of making important decisions; for the former the critical decisions they have to make concerning the selection of the firms that will receive financial support or subsidies for digital investment, employment of ICT personnel or digital operating expenses in general (enabling the focus of the scarce financial resources on the most digitally vulnerable firms, which are in greatest need of government support); for the latter the critical decisions they have to make concerning the selection of the firms they will grant loans to, or firms they will invest in (enabling them to grant loans to and invest in firms that will not show high digital vulnerability to economic crisis, and therefore are not at risk of digital obsolescence ,which can have negative consequences for their future competitiveness or even for their survival). Furthermore, our methodology provides a useful tool for ICT firms, as well as consulting firms offering ICT-related services, for enhancing/augmenting their process of making important decisions concerning the selection of firms they will focus on during economic crises (enabling them to focus their marketing and sales activities on firms that will not show high digital vulnerability to economic crisis, thus increasing the likelihood of substantial sales revenue).

However, further research is required in the following directions: a) further application of the proposed methodology, using larger datasets, with more dependent variables (i.e. more types of digital investment as well as digital operating expenses) and more independent variables (i.e. more firms’ characteristics), in various national and sectoral contexts (experiencing different types and intensities of economic crises), and even for different categories of investments and operating expenses (beyond the digital ones); b) investigation of the use of other pre-processing algorithms (e.g., oversampling and class-balancing algorithms) as well as AI/ML algorithms (e.g., deep learning ones) for achieving higher prediction performance; c) investigation of the combination of OGD with other sources of data about firms (e.g. from other government agencies, from private firms, such as private business information databases), in order to obtain more information about firms that might improve the performance of the prediction of their pattern of vulnerability to economic crises; d) investigation of the large-scale application of the proposed methodology, and the main challenges it may pose (possibly using as a starting point the challenges of using OGD for training AI models mentioned in [24]).

## ACKNOWLEDGMENTS

This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 955569.

## References

[1] OECD, Responding to the Economic Crisis Fostering Industrial Restructuring and Renewal. OECD Publishing, (2009a).

[2] Keeley, B, Love, P.: From Crisis to Recovery - The Causes, Course and Consequences of the Great Recession. OECD Publishing, Paris, (2012).

[3] Knoop, T. A.: Recessions and Depressions: Understanding Business Cycles - 2nd edition. Praeger Santa Barbara, California, (2015).

[4] Allen, R. E.: Financial Crises and Recession in the Global Economy - Fourth Edition. Edward Elgar Publications, Cheltenham, UK, (2017).

[5] Paris.S. Saifi and J. Horowitz, "Blackouts and soaring prices: Pakistan's economy is on the brink." https://edition.cnn.com/2023/02/02/economy/pakistan-economy-crisis/index.html

[6] E. Kottika *et al.*, "We survived this! What managers could learn from SMEs who successfully navigated the Greek economic crisis," *Industrial Marketing Management*, vol. 88, pp. 352–365, Jul. 2020, doi: 10.1016/j.indmarman.2020.05.021.

[7] Loukis, E., Arvanitis, S., Myrtidis, D.: ICT-Related Behaviour of Greek Banks in the Economic Crisis. Information Systems Management, 38(1), 79-91 (2021).

[8] Baldwin, R. and Di Mauro, B. W.: Mitigating the COVID economic crisis: Act fast and do whatever it takes. Center of Economic Policy Research Press, London, (2020).

[9] L. Oliveras *et al.*, "Energy poverty and health: Trends in the European Union before and during the economic crisis, 2007–2016," *Health & Place*, vol. 67, p. 102294, Jan. 2021, doi: 10.1016/j.healthplace.2020.102294.

[10] A. Dagoumas and F. Kitsios, "Assessing the impact of the economic crisis on energy poverty in Greece," *Sustainable Cities and Society*, vol. 13, pp. 267–278, Oct. 2014, doi: 10.1016/j.scs.2014.02.004.

[11] D. Stylidis and M. Terzidou, "Tourism and the economic crisis in Kavala, Greece," *Annals of Tourism Research*, vol. 44, pp. 210–226, Jan. 2014, doi: 10.1016/j.annals.2013.10.004.

[12] M. Das, "Economic crisis in Sri Lanka causing cancer drug shortage," *The Lancet*

*Oncology*, vol. 23, no. 6, p. 710, Jun. 2022, doi: 10.1016/S1470-2045(22)00254-6.
[13] Arvanitis, S., Loukis, E.: Factors Factors explaining ICT investment behavior of firms during the 2008 economic crisis. *Information Systems Management* (in-press), (2023).
[14] Khatiwada, S.: Stimulus Packages to Counter Global Economic Crisis -– A review. International Institute for Labour Studies, Geneva, (2009).
[15] Coenen, G., Straub, R., Trabandt, M.: Gauging the Effects of Fiscal Stimulus Packages in the Euro Area. Working Paper 1483, European Central Bank, Frankfurt am Main, Germany, (2012).
[16] Kalinowski, T.: Crisis management and the diversity of capitalism: fiscal stimulus packages and the East Asian (neo-)developmental state. *Economy and Society* 44(2), 244-270 (2015).
[17] Taylor, J.: Fiscal Stimulus Programs During the Great Recession. Economics Working Paper 18117, Hoover Institution, Stanford, CA, (2018).
[18] Witten, I. H., Frank, E., Hall, M. A., Pal, C. J.`: Data mining - practical machine learning tools and techniques. Morgan Kaufmann, Amsterdam, London, (2017).
[19] Blum, A., Hopcroft, J., Kannan, R.: Foundations of Data Science. Cambridge University Press, Cambridge, (2020).
[20] Russell, S., Norvig, P.: Artificial Intelligence. A Modern Approach -– 3rd Edition. Pearson, Essex, UK, (2020).
[21] Y. Charalabidis, A. Zuiderwijk, C. Alexopoulos, M. Janssen, T. Lampoltshammer, E. Ferro. The World of Open Data - Concepts, Methods, Tools and Experiences. Springer International Publishing AG, Switzerland, (2018).
[22] M. Ali, C. Alexopoulos, and Y. Charalabidis, (2022). “A Comprehensive Review of Open Data Platforms, Prevalent Technologies, and Functionalities,” in Proceedings of the 15th International Conference on Theory and Practice of Electronic Governance, New York, NY, USA, pp. 203–214. doi: 10.1145/3560107.3560142.
[23] Y. Gao, M. Janssen and C. Zhang.: "Understanding the evolution of open government data research: towards open data sustainability and smartness", *International Review of Administrative Sciences* 89(1), 59-75 (2013).
[24] UNESCO, Open data for AI, United Nations Educational, Scientific and Cultural Organization, (2023)
[25] DeSousa, W. G., DeMelo, E. R. P., De Souza Bermejo, P. H., Sous Farias, R. A., 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), 101392 (2019).
[26] Zuiderwijk, A., Chen, Y. C., Salem, F.: "Implications of the use of artificial intelligence in public governance: A systematic literature review and a research agenda". Government Information Quarterly 38(3), 101577 (2021).
[27] Medaglia, R., Gil-Garcia, R. and Pardo, T. A.. "Artificial Intelligence in Government: Taking Stock and Moving Forward". Social Science Computer Review (in-press), (2021).
[28] Manzoni, M., Medaglia, R., Tangi, L., Van Noordt, C., Vaccari, L. and Gattwinkel, D.. "AI Watch. Road to the adoption of Artificial Intelligence by the public sector". Publications Office of the European Union, Luxembourg, (2022).
[29] Van Noordt, C., Misuraca, G.. "Artificial intelligence for the public sector: results of landscaping the use of AI in government across the European Union". Government Information Quarterly (in-press), (2022).
[30] R. Madan , M. Ashok. "AI adoption and diffusion in public administration: A systematic

literature review and future research agenda". Government Information Quarterly 40(1) 101774, (2023).
[31] Leavitt, H. J.. "Applied organizational change in industry: Structural, technological and humanistic approaches". In: J. G. March (Ed.), Handbook of organizations -– 3rd ed., 1144–1170. Rand McNally & Company, Chicago, IL, (1970).
[32] Scott-Morton, M.S.. "The Corporation of the 1990s" Oxford University Press, New York, (1991).
[33] Whittington, R., Regner, P., Angwin, D., Johnson, G., Scholes, K.: "Exploring Strategy - 12th edn". Pearson Education Limited, Harlow, UK, (2020).
[34] M. Dudjak and G. Martinović, “In-Depth Performance Analysis of SMOTE-Based Oversampling Algorithms in Binary Classification,” Int. j. electr. comput. eng. syst. (Online), vol. 11, no. 1, pp. 13–23, Apr. 2020, doi: 10.32985/ijeces.11.1.2.

## **Appendix -** Dependent Variables' Questions

How important was the impact of the economic crisis during the period 2009-2014 on the following categories of ICT related investment and operating expenses in your company?

| | Increase | Negligible Impact | Small Decrease | Moderate Decrease | Large Decrease | Very Large Decrease |
|---|---|---|---|---|---|---|
| Total hard and soft ICT investments (in ICT hardware, ICT software, ICT training and ICT consulting) | | | | | | |
| Investments in ICT hardware | | | | | | |
| Investments in ICT software | | | | | | |
| Investments in ICT training of your staff | | | | | | |
| Investment in ICT consulting | | | | | | |
| Total ICT operating expenses (ICT personnel and services) | | | | | | |
| ICT staff expenses | | | | | | |
| ICT outsourcing expenses | | | | | | |
| Cloud Computing services expenses | | | | | | |