A Machine Learning Based Solution for Forecasted Economics Predicting Business Dynamics Across Europe Using Open Government Data

Mustafa I Al-Karkhi[1], Grzegorz Rzadkowski[2]

[1]Mechanical Engineering Department, University of Technology, Baghdad, Iraq
[2]Department of Finance and Risk Management, Warsaw University of Technology, Narbutta 85, 02-524 Warsaw, Poland

Abstract

In an era where data-driven decision-making is paramount, this paper explores the use of advanced statistical methods and machine learning to forecast business dynamics within a specific continent, Europe. The importance of forecasted economics is discussed in the introduction, followed by a review of the relevant literature which highlights recent work and forecasts. The work has leveraged Open Government Data (OGD) from Scotland and Wales for the period 2010 to 2023, this study also employs an Artificial Neural Network (ANN) for regression analysis to predict the growth of businesses. Detailed attention is given to the parameters of the ANN used to ensure methodological transparency. The model's effectiveness is evaluated using Mean Absolute Percent-age Error (MAPE) and the coefficient of determination (R^2), with remarkable results presented as 0.8% and 0.97, respectively. Later, results are visually represented through various techniques to the purpose of comparing predicted outcomes with actual data. The paper concludes by outlining the significant contributions of the study and it also emphasizes the enhanced capability of ANNs in economic forecasting and their potential impact on policy-making.

Keywords

Forecasted Economics, Artificial Neural Networks, Open Government Data, Business Dynamics

1. Introduction

Over the past few decades, European economics and businesses have experienced significant shifts marked by periods of growth and stagnation [1, 2, 3]. These changes are evident in the fluctuating number of businesses across the continent, a key indicator of economic health and future conditions. In regions like Scotland and Wales, the economic landscape has been notably influenced by both regional policies and broader European trends [4, 5]. These areas have seen a rise in businesses, driven by technological ad-vancements, changing consumer behaviors, and government support through various initiatives. For policymakers and business leaders, understanding these trends is essential, as the economic environment impacts both macroeconomic stability and microeconomic activities [6, 7, 8]. An increase in businesses usually indicates a strong economy but can also result in greater competition and market saturation. Analyzing these patterns through forecasted economics provides insight into the forces shaping these trends. Focusing on Europe, especially Scotland and Wales, allows for a unique study of these phenomena on both local and continental levels. This analysis helps compare dif-ferent regions and understand the specific economic drivers and challenges faced by businesses. Thus, examining business dynamics in Europe is crucial for predicting fu-ture economic health and planning strategic policies that promote sustainable growth. The significance of data-driven forecasting in economic decision-making has been increasingly recognized in recent scholarly work. Rodgers et al. (2024) introduced a me-thodical approach known as the Business-Driven Data-Supported (BDDS) process, which emphasized the importance of starting with clearly defined business problems to effectively harness Big Data and AI for actionable insights [9]. Their process is well-supported by a prescriptive framework, the Da-ta-to-Information-Extraction-Methodology (DIEM), which has demonstrated efficacy across various industries, particularly highlighted through a healthcare industry exam-ple. Moreover, in terms of urban economics, Hatami et al. (2024) addressed the gap in economic forecasting by the development of a spatiotemporal deep learning model that predicts the performance of non-business services in small urban areas [10, 11] or for peer to peer communications and video streaming [12, 3, 13] and other applications [14, 15]. The re-search utilized LSTM networks and underscored the complexities of economic forecasting at the micro-geographic level. The findings revealed significant predictive accuracy in employment, business sales volume, and labor productivity within Mecklenburg County, North Carolina. Further emphasizing the theoretical underpinnings of business dynamics, Dominko et al. (2023) conducted a bibliometric analysis focusing on the circular economy within business and economics fields [16]. The study highlighted the need for action-oriented

SYSYEM 2024: 10th Scholar's Yearly Symposium of Technology, Engineering and Mathematics, Rome, December 2-5, 2024
Mustafa.I.Alkarkhi@uotechnology.edu.iq (M. I. Al-Karkhi); Grzegorz.rzadkowski@pw.edu.pl (G. Rzadkowski)

research to expedite the transition from linear to circular economic models which sug-gests a focus on sustainable supply chains, waste management, and business model in-novation. Furthermore, machine learning's role in forecasting and estimation was elabo-rated by Ahamed et al. (2023) as they evaluated various machine learning algorithms for their applicability in forecasting sales of truck components [17]. The comparative analy-sis presented the superiority of the Random Forest Regression model over others like Simple Linear Regression and Ridge Regression, thus providing a robust foundation for future business forecasting endeavors. The dynamics of economic growth forecasts have also been explored in relation to business cycles, also useful for integration of different communication networks [18, 19]. Huh and Kim (2020) investigated revisions in growth forecasts and uti-lized data from the Survey of Professional Forecasters [20]. The study found a distinct asymmetry in forecast revisions across economic cycles, with more significant adjust-ments during economic contractions compared to expansions. Lastly, Aminullah (2024) delved into forecasting technology innovation and economic growth in Indonesia by the application of a feedback economics perspective [21, 19]. The study emphasized the inte-gration of general-purpose technologies with industrial policies to ensure sustainable and inclusive growth post-COVID-19. The study also provided a unique insight into the reindustrialization process and its implications for economic policy. Although extensive research has been conducted on economic forecasting with ma-chine learning, gaps persist, particularly in applying these technologies to predict busi-ness dynamics using open government data in specific European regions. Previous studies have mainly focused on broader economic indicators or non-business services, often overlooking the potential of neural networks to utilize publicly available datasets for predicting business growth. Furthermore, the unique economic contexts of Scotland and Wales have seldom been the subject of such detailed predictive analyses. This study addresses these gaps by employing an advanced Artificial Neural Network (ANN) to analyze open government data from 2010 to 2023, aiming to forecast business growth in Scotland and Wales. This approach not only deepens the understanding of regional economic dynamics but also advances the literature by providing methodological innovations in using neural networks for economic forecasting. The novelty of this work lies in its integration of detailed neural network parameterization with a focused application on open government data, offering pivotal insights for policymakers and economic strategists. The remainder of the paper is structured as follows: Section 2 describes the data collection process and the setup of the neural network model. Section 3 delves into the specifics of the ANN configuration and the rationale behind the chosen parameters. Sec-tion 4 presents the findings from the model. Finally, Section 5 summarizes the key con-tributions of the study and suggests avenues for further research.

2. Experimental Methodology: Design, Materials and Methods

Open Government Data (OGD) is a central data management approach used in the transparency, accountability, and public engagement spectrum to facilitate the release of data through official governmental portals. The primary purpose of these portals is not only to be a data repository, but also a platform that enables various stakeholders to discover and reuse the important information. OGD is used in a growing number of re-search activities such as trend analysis, predicting economic indicators, and evaluating policy impacts because it is structured and contains a huge amount of data. In the context of exploring the dynamics of business growth within Scotland and Wales, the methodology adopted in a recent study provides a relevant framework [22, 23, 24, 25, 26]. It was reported that data were systematically collected from the open govern-ment portals of 28 European Union countries with a focus on various parameters and indicators over a three-year period. Likewise, but not similarly, the current study had used the same OGD but for a period from 2010 till 2023, for an updated timeline. This comprehensive dataset was smoothly assembled through manual retrieval from public sources and official OGD portals, which are freely accessible on the web. Furthermore, to synthesize the data, density-based spatial clustering techniques (DBSCAN) were ap-plied, and the cluster validity was assessed using the Davies–Bouldin index as depicted in Figure 1. Drawing from the same sources, this study will utilize the dataset compiled by the aforementioned references and then to apply it to forecast business growth through the deployment of ANN models. The approach ensures a robust analytical ba-sis for the examination of regional economic trends within the specified timeframe.

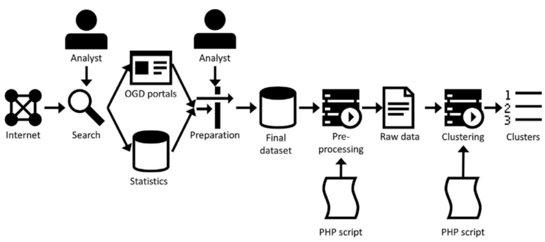

Figure 1: Dataset creation process [22]

3. Machine Learning

Machine learning (extended from artificial intelligence) is centered around the idea that we can make computers learn based on data and make decisions without writing a specific code to accomplish this in each case [27, 28, 29, 30]. It automatically detects patterns, creates predictions, and extracts actionable insights from large datasets, changing research across various disciplines [31, 32, 33, 34, 35]. At the heart of machine learning are algorithms that improve automatically through experience and by processing data. Of these algorithms, Artificial Neural Networks (ANNs) participate by imitating the neural structure of the human brain and have the ability to estimate non-linear relationships in data [36, 37, 38, 39, 40, 41]. ANNS are modeled lightly based on the human brain; they are composed of connected nodes, or 'neurons.' The nodes are made of layers that process the input data by various transformations. Hierarchy of Networks - These networks have an input layer that takes in the data, some hidden layers that process the data, and at the end, an output layer that gives the results of the prediction or, in some cases, the classification. Feedforward algorithms handle large numbers of datasets having complexity, which makes them perfect for image recognition, natural language processing, and economic forecasting. In economic terms, this provides vital information in order to predict the future, which is aimed at influencing the sort of strategies that organizations use when planning for the future In the implementation of the ANN for this study, key parameters and their specifications are important for functionality and replication, as depicted in Figure 2 and detailed in Table 1. Figure 2 illustrates the basic structure of the ANN, where y denotes the output,w_i represents the weights associated with the inputs, and b indicates the bias, all processed through an activation function denoted by f This schematic provides a visual representation of the data flow within the network. Table 1 outlines the specific parameters used in the construction of the ANN. The network comprises two hidden layers, the first one contains ten neurons and the second contains four. The Adam solver was chosen to optimize the network since it is known for its efficiency in handling large datasets and noisy problems. The training process was conducted over 500 iterations which allowed sufficient time for the network to converge and adjust its weights and biases to minimize prediction errors to maintain robustness in the forecasting model.

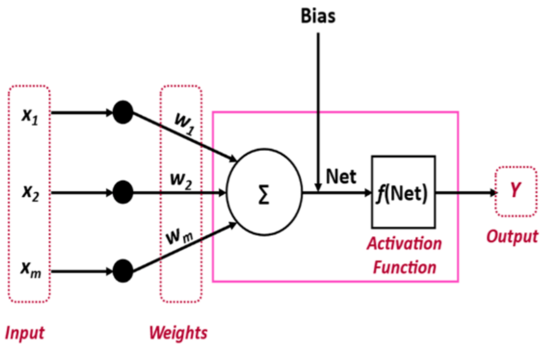

Figure 2: Neural network workflow [38]

Table 1
ANN used parameters

Parameter	Value
Number of hidden layers	2
Number of neuron in first hidden layer	10
Number of neuron in second hidden layer	4
Solver	Adam
Number of iterations	500

4. Results and Discussion

4.1. Data Visualization

In the exploration of business growth dynamics within the United Kingdom, partic-ularly in Scotland and Wales, visual representations provide a clear depiction of trends over the period from 2010 to 2023. Figure 3 presents a free visualization of the number of businesses in Scotland as it illustrates a general trend of growth with some fluctua-tions over the examined years. Notably, a peak is observed in the years leading up to 2019, followed by a noticeable decline, which may correlate with economic or policy changes during that period, or the pandemic as we all know it. Conversely, Figure 4 fo-cuses on Wales by the showcasing of a slightly different trajectory. The visualization in-dicates a steady increase in the number of businesses up to 2020, followed by a dip likely influenced by external economic factors, before a modest recovery in subsequent years. These visual analyses highlight the regional economic conditions and also pro-vide a comparative perspective on how different areas within Europe are evolving in terms of business establishment and growth. The data offer a general valuable insight for regional economic planning and development strategies.

4.2. Forecasts and Regression Results

In evaluating the predictive capabilities of the neural network model applied to the business dynamics in Scotland and Wales, the forecasted results are quantitatively assessed through key performance indicators, namely the MAPE and R^2. As detailed in Table 2, the model achieved a MAPE of 1.1% and an R^2 of 0.966 for Scotland, this is

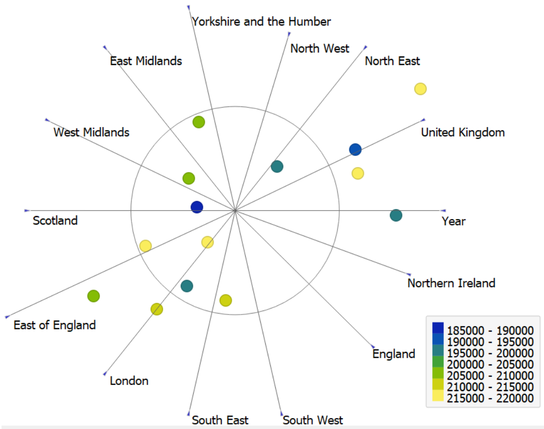

Figure 3: Free visualization for the number of business in Scotland for a period of [2010-2023]

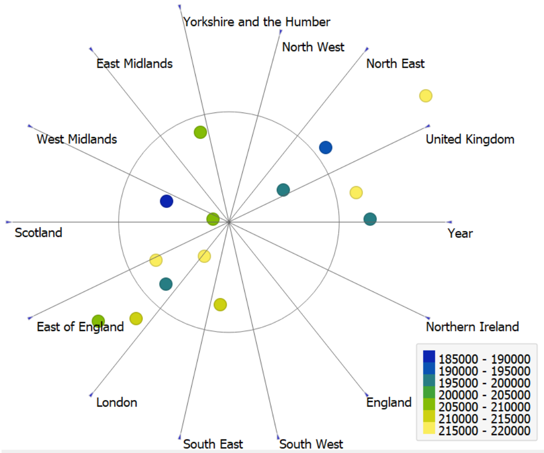

Figure 4: Free visualization for the number of business in Wales for a period of [2010-2023]

an indication of a high degree of accuracy and reliability in the forecasts. Similarly, for Wales, the model demonstrated even more precise predictions with a MAPE of 0.8% and an R^2 of 0.97 which suggests that the forecasts are nearly exact replicas of the observed data.

The regression analyses that corroborate these practical implications are illustrated in Figures 5 and 6. In Figure 5 the regression line is given to help determine the forecasted versus non-forecasted number of businesses for Scotland, such as a forecasted confidence at 291,996 catching up on an original count from 2010 of 285,000, and a forecast of (303,065) approaching near the original value of 298,300 in 2023. The close alignment between the predictions of the current approach and the future results provided by the administrative data model underlines the model's ability to capture the genuine underlying trends in the Scottish business population. Similarly, Figure 6 is for Wales and shows that the forecasted line of best fit is closely aligned to the actual business count, meaning the initial forecast of around 193,571 closely mirrors the actual 190,800 business count in 2010 and expected 217,067 by 2023 is therefore not far from the actual 219,000 businesses. Not only does this visual representation exacerbate the statistical results, it also clearly shows just how well our model was able to adapt to the economic circumstances in Wales. The low MAPE figures and high R^2 scores for these results confirm the ANN based technique to be a powerful tool for business number forecasting. It is concluded that the results of the model show that it has the potential to be an important tool when it comes to the economic analysis and policy making in regions with dynamic economic activity.

Table 2
Forecasted results

Forecasted Business Number of:	MAPE(%)	R^2
Scotland	1.1	0.966
Wales	0.8	0.97

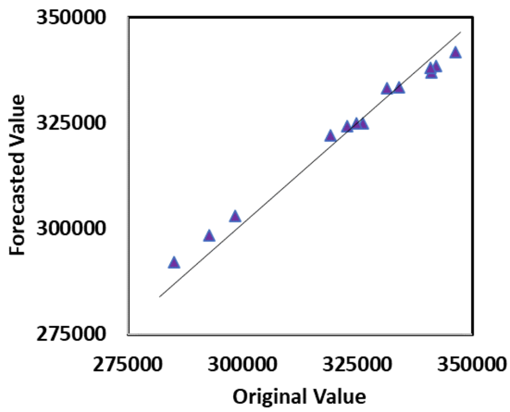

Figure 5: Regression line for the forecasts of number of businesses in Scotland

5. Conclusions and Future Directions

This study successfully employed an advanced ANN to forecast the number of businesses in Scotland and Wales from 2010 to 2023, utilizing OGD. The ANN model demonstrated high predictive accuracy, as evidenced by MAPE and the R^2 values, which were notably low and high respectively, indicating the model's efficacy. The re-search significantly contributes to the field by highlighting the

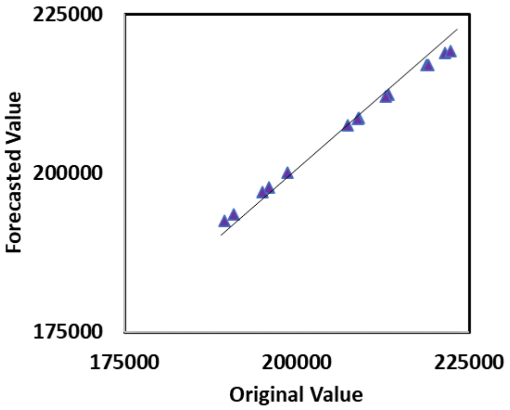

Figure 6: Regression line for the forecasts of number of businesses in Wales

applicability of neural networks in economic forecasting and providing a reliable methodological framework for regional economic analysis. Furthermore, it enriches the understanding of business dynamics in the specific context of Scotland and Wales. This had offered insights that can assist policymakers and economic developers in strategic planning and economic assessment. Looking ahead, the study opens several avenues for future research. Firstly, ex-panding the geographical scope to include more regions or comparing different eco-nomic environments could further validate and refine the forecasting model. Addition-ally, incorporating other predictive variables such as economic policies, market trends, and socio-economic factors could enhance the model's comprehensiveness and accuracy. Future work could also explore the integration of real-time data feeds to enable dynam-ic forecasting that adapts to new data as it becomes available. Lastly, applying different machine learning techniques and comparing their performance could provide deeper insights into the most effective methods for economic forecasting. This would poten-tially lead to the more robustness in predictive models for various

References

[1] D. W. Hands, The many faces of unification and pluralism in economics: The case of paul samuelson's foundations of economic analysis, Studies in History and Philosophy of Science 88 (2021) 209–219. URL: https://www.sciencedirect.com/science/article/pii/S0039368121000911. doi:https://doi.org/10.1016/j.shpsa.2021.06.008.

[2] S. Gunessee, T. Lane, Changing perceptions about experimentation in economics: 50 years of evidence from principles textbooks, Journal of Behavioral and Experimental Economics 107 (2023) 102086. URL: https://www.sciencedirect.com/science/article/pii/S221480432300112X. doi:https://doi.org/10.1016/j.socec.2023.102086.

[3] G. Ciccarella, F. Vatalaro, A. Vizzarri, Content delivery on ip network: Service providers and tv broadcasters business repositioning, in: 2019 3rd International Conference on Recent Advances in Signal Processing, Telecommunications & Computing (SigTelCom), 2019, pp. 149–154. doi:10.1109/SIGTELCOM.2019.8696155.

[4] R. T. Edwards, J. Charles, S. Thomas, H. Howson, P. Bradley, Use of health economics and programme budgeting and marginal analysis at a national level to prioritise public health improvement programmes in wales, The Lancet 380 (2012) S75.

[5] M. I. Al-Karkhi, Statistical analysis and open innovation in economic growth of scottish business smes for sustainable development, Journal of Open Innovation: Technology, Market, and Complexity 10 (2024) 100275.

[6] G. M. Martin, D. T. Frazier, W. Maneesoonthorn, R. Loaiza-Maya, F. Huber, G. Koop, J. Maheu, D. Nibbering, A. Panagiotelis, Bayesian forecasting in economics and finance: a modern review, International Journal of Forecasting (2023).

[7] Q. Liang, J. Huang, M. Liang, J. Li, Economic growth targets and bank risk exposure: Evidence from china, Economic Modelling (2024) 106702.

[8] Z. Deng, S. Song, N. Jiang, R. Pang, Sustainable development in china? a nonparametric decomposition of economic growth, China Economic Review 81 (2023) 102041.

[9] M. Rodgers, S. Mukherjee, B. Melamed, A. Baveja, A. Kapoor, Solving business problems: the business-driven data-supported process, Annals of Operations Research (2024) 1–37.

[10] F. Hatami, A. P. Poor, J.-C. Thill, Non-business services performance forecasting for small urban areas using a spatiotemporal deep learning model, Cities 152 (2024) 105141.

[11] C. Napoli, G. Pappalardo, E. Tramontana, R. K. Nowicki, J. T. Starczewski, M. Woźniak, Toward work groups classification based on probabilistic neural network approach, in: Lecture Notes in Artificial Intelligence (Subseries of Lecture Notes in Computer Science), volume 9119, 2015, p. 79 – 89. doi:10.1007/978-3-319-19324-3_8.

[12] C. Napoli, G. Pappalardo, E. Tramontana, Improving files availability for bittorrent using a diffusion model, in: Proceedings of the Workshop on Enabling Technologies: Infrastructure for Collaborative Enterprises, WETICE, 2014, p. 191 – 196. doi:10.1109/WETICE.2014.65.

[13] G. Capizzi, C. Napoli, L. Paternò, An innovative

hybrid neuro-wavelet method for reconstruction of missing data in astronomical photometric surveys, in: Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), volume 7267 LNAI, 2012, p. 21 – 29. doi:10.1007/978-3-642-29347-4_3.

[14] F. Fiani, S. Russo, C. Napoli, An advanced solution based on machine learning for remote emdr therapy, Technologies 11 (2023). doi:10.3390/technologies11060172.

[15] S. Falciglia, F. Betello, S. Russo, C. Napoli, Learning visual stimulus-evoked eeg manifold for neural image classification, Neurocomputing 588 (2024). doi:10.1016/j.neucom.2024.127654.

[16] M. Dominko, K. Primc, R. Slabe-Erker, B. Kalar, A bibliometric analysis of circular economy in the fields of business and economics: towards more action-oriented research, Environment, Development and Sustainability 25 (2023) 5797–5830.

[17] S. F. Ahamed, A. Vijayasankar, M. Thenmozhi, S. Rajendar, P. Bindu, T. S. M. Rao, Machine learning models for forecasting and estimation of business operations, The Journal of High Technology Management Research 34 (2023) 100455.

[18] R. Giuliano, F. Mazzenga, E. Innocenti, A. Vizzarri, Integration of video and radio technologies for social distancing, IEEE Communications Magazine 59(9) (2021) 30–35.

[19] C. Napoli, G. Pappalardo, E. Tramontana, An agent-driven semantical identifier using radial basis neural networks and reinforcement learning, in: CEUR Workshop Proceedings, volume 1260, 2014.

[20] S. Huh, I. Kim, Growth forecast revisions over business cycles: Evidence from the survey of professional forecasters, Economics letters 196 (2020) 109541.

[21] E. Aminullah, Forecasting of technology innovation and economic growth in indonesia, Technological Forecasting and Social Change 202 (2024) 123333.

[22] S. de Juana-Espinosa, S. Luján-Mora, Open government data portals in the european union: A dataset from 2015 to 2017, Data in brief 29 (2020) 105156.

[23] S. de Juana-Espinosa, S. Luján-Mora, Open government data portals in the european union: Considerations, development, and expectations, Technological Forecasting and Social Change 149 (2019) 119769.

[24] C. Napoli, C. Napoli, V. Ponzi, A. Puglisi, S. Russo, I. E. Tibermacine, Exploiting robots as healthcare resources for epidemics management and support caregivers, in: CEUR Workshop Proceedings, volume 3686, 2024, p. 1 – 10.

[25] V. Marcotrigiano, G. D. Stingi, S. Fregnan, P. Magarelli, P. Pasquale, S. Russo, G. B. Orsi, M. T. Montagna, C. Napoli, C. Napoli, An integrated control plan in primary schools: Results of a field investigation on nutritional and hygienic features in the apulia region (southern italy), Nutrients 13 (2021). doi:10.3390/nu13093006.

[26] V. Ponzi, S. Russo, A. Wajda, R. Brociek, C. Napoli, Analysis pre and post covid-19 pandemic rorschach test data of using em algorithms and gmm models, in: CEUR Workshop Proceedings, volume 3360, 2022, p. 55 – 63.

[27] L. A. Al-Haddad, et al., Enhancing wind tunnel computational simulations of finite element analysis using machine learning-based algorithms, Eng Technol J 42 (2023) 1–9.

[28] L. A. Al-Haddad, A. A. Jaber, Improved uav blade unbalance prediction based on machine learning and relieff supreme feature ranking method, Journal of the Brazilian Society of Mechanical Sciences and Engineering 45 (2023) 463.

[29] L. A. Al-Haddad, A. A. Jaber, S. A. Al-Haddad, Y. M. Al-Muslim, Fault diagnosis of actuator damage in uavs using embedded recorded data and stacked machine learning models, The Journal of Supercomputing 80 (2024) 3005–3024.

[30] L. A. Al-Haddad, A. Jaber, Applications of machine learning techniques for fault diagnosis of uavs (2022).

[31] G. C. Cardarilli, L. Di Nunzio, R. Fazzolari, D. Giardino, M. Re, A. Ricci, S. Spano, An fpga-based multi-agent reinforcement learning timing synchronizer, Computers and Electrical Engineering 99 (2022) 107749.

[32] G. C. Cardarilli, L. Di Nunzio, R. Fazzolari, D. Giardino, M. Matta, M. Patetta, M. Re, S. Spanò, Approximated computing for low power neural networks, Telkomnika (Telecommunication Computing Electronics and Control) 17 (2019) 1236–1241.

[33] G. Lo Sciuto, G. Susi, G. Cammarata, G. Capizzi, A spiking neural network-based model for anaerobic digestion process, in: 2016 International Symposium on Power Electronics, Electrical Drives, Automation and Motion (SPEEDAM), IEEE, 2016, pp. 996–1003.

[34] F. Bonanno, G. Capizzi, G. Lo Sciuto, C. Napoli, G. Pappalardo, E. Tramontana, A cascade neural network architecture investigating surface plasmon polaritons propagation for thin metals in openmp, in: Artificial Intelligence and Soft Computing: 13th International Conference, ICAISC 2014, Zakopane, Poland, June 1-5, 2014, Proceedings, Part I 13, Springer, 2014, pp. 22–33.

[35] G. Lo Sciuto, G. Capizzi, R. Shikler, C. Napoli, Organic solar cells defects classification by using a new feature extraction algorithm and an ebnn with an innovative pruning algorithm, International

Journal of Intelligent Systems 36 (2021) 2443–2464.
[36] L. A. Al-Haddad, W. H. Alawee, A. Basem, Advancing task recognition towards artificial limbs control with relieff-based deep neural network extreme learning, Computers in Biology and Medicine 169 (2024) 107894.
[37] M. Y. Fattah, L. A. Al-Haddad, M. Ayasrah, A. A. Jaber, S. A. Al-Haddad, Coupled finite element and artificial neural network analysis of interfering strip footings in saturated cohesive soils, Transportation Infrastructure Geotechnology (2024) 1–18.
[38] L. A. Al-Haddad, A. A. Jaber, An intelligent fault diagnosis approach for multirotor uavs based on deep neural network of multi-resolution transform features, Drones 7 (2023) 82.
[39] L. A. Al-Haddad, L. Ibraheem, A. I. EL-Seesy, A. A. Jaber, S. A. Al-Haddad, R. Khosrozadeh, Thermal heat flux distribution prediction in an electrical vehicle battery cell using finite element analysis and neural network, Green Energy and Intelligent Transportation (2024) 100155.
[40] L. A. Al-Haddad, Y. M. Al-Muslim, A. S. Hammood, A. A. Al-Zubaidi, A. M. Khalil, Y. Ibraheem, H. J. Imran, M. Y. Fattah, M. F. Alawami, A. M. Abdul-Ghani, Enhancing building sustainability through aerodynamic shading devices: an integrated design methodology using finite element analysis and optimized neural networks, Asian Journal of Civil Engineering (2024) 1–14.
[41] D. W. Hands, The many faces of unification and pluralism in economics: The case of paul samuelson's foundations of economic analysis, Studies in History and Philosophy of Science 88 (2021) 209–219.