Idea: Future Research
Volume 1, Issue 1, March 2023
ISSN 2987-1565 (Online) | ISSN 2987-2715 (Print)
Published by:
Ideas for Future Research and Technology

Integrating the neural network into the stochastic DEA model

Hengki Tamando Sihotang[1], J. Lavemaau[2], Fristi Riandari[3], Firta Sari Panjaitan[4], Sonya Enjelina Gorat[5], Juliana Batubara[6]
[1,3,4,5,6]Departemen Riset, Institute of Computer Science, Indonesia
[2]International Office, Institute of Computer Science, Ceko

Introduction

Data Envelopment Analysis (DEA) is a popular linear programming technique used for measuring the relative efficiency of decision-making units (DMUs) operating under multiple inputs and outputs (Angiz et al., 2013)(Joro et al., 1998). It has found widespread use in many fields, including finance, healthcare, and manufacturing (Esmaeilzadeh, 2020). However, one of the challenges of using DEA is that it assumes that the input and output data of the DMUs are deterministic. In reality, these data are often subject to uncertainty and noise (Khaki et al., 2012)(Amirkhan et al., 2018)(Wu et al., 2022)(Amirteimoori et al., 2023).

Stochastic DEA is a variant of DEA that incorporates uncertainty in the input and output data by assuming that these data are known probability distributions (Sengupta, 1987)(Dotoli et al., 2016)(Aydın & Toklu, 2023)(Amirteimoori et al., 2023). This approach can produce more robust efficiency scores by accounting for the uncertainty in the data (Habib et al., 2022)(Qu et al., 2022). However, stochastic DEA models still rely on assumptions about the probability distributions of the input and output variables, which may not always hold true in practice (Labijak-Kowalska & Kadziński, 2023).

To address these challenges, researchers have proposed integrating neural networks (NN) into stochastic DEA models (Modhej & Dahimavi, 2022)(Maniati et al., 2022)(Nazari-Shirkouhi et al., 2023)(Shi & Zhao, 2023). NNs are a class of machine learning models that are capable of learning complex patterns in data and making predictions based on those patterns (Brumm et al., 2023)(Bouzidi et al., 2022). By incorporating NNs into stochastic DEA models, it is possible to improve the accuracy of efficiency measurement and prediction while reducing computational complexity (Tang et al., 2022)(Zhu et al., 2021).

The basic idea behind integrating NNs into stochastic DEA models is to use the NN to estimate the distribution of the input and output data of the DMUs, and then use the stochastic DEA model to calculate the efficiency scores of the DMUs based on these estimated distributions (Ang et al., 2021)(Gholizadeh et al., 2022)(Namakin et al., 2021)(Cinaroglu, 2023)(Zarrin & Brunner, 2023). The NN can be trained on historical

Corresponding Author: Dr. Hengki Tamando Sihotang, M. Kom (hengki_tamando@iocscience.org)

data to learn the relationships between the input and output variables, and can then be used to estimate the probability distributions of these variables for new DMUs (Balak et al., 2021)(Aggarwal et al., 2021)(Yazdanparast et al., 2021).

Several studies have explored the integration of NNs into stochastic DEA models. For example, Lu et al. (2009) proposed a method that uses a feedforward NN to estimate the conditional mean and variance of the input and output data of the DMUs, and then uses a stochastic DEA model to calculate the efficiency scores (Boubaker et al., 2023)(Wang et al., 2022)(Kainthura & Sharma, 2022). Zhang et al. (2014) proposed a similar method that uses a radial basis function NN to estimate the conditional distribution of the input and output data of the DMUs (Pendharkar, 2023).

Other researchers have proposed more advanced NN architectures, such as deep neural networks (DNNs) and convolutional neural networks (CNNs), for integrating into stochastic DEA models. For example, Wang et al. (2020) proposed a method that uses a DNN to estimate the probability distributions of the input and output data of the DMUs, and then uses a stochastic DEA model to calculate the efficiency scores (Dalei & Joshi, 2023). Chen et al. (2021) proposed a method that uses a CNN to extract features from the input and output data of the DMUs, and then uses a stochastic DEA model to calculate the efficiency scores based on these features (Damayanti et al., 2023).

The integration of NNs into stochastic DEA models has several advantages (Fallahpour et al., 2016). First, it can improve the accuracy of efficiency measurement by better capturing the underlying distribution of the input and output data of the DMUs. Second, it can reduce the computational complexity of the stochastic DEA model by reducing the number of Monte Carlo simulations required to estimate the efficiency scores. Third, it can handle non-linear relationships between the input and output data of the DMUs, which cannot be captured by traditional linear models.

However, there are also some challenges associated with integrating NNs into stochastic DEA models. One of the main challenges is selecting an appropriate NN architecture and training algorithm for a given problem. Different NN architectures may perform better on different types of data, and selecting the wrong architecture could result in poor performance. Additionally, NNs require a large amount of training data to learn the underlying patterns in the data, which may not always be available in practice (Apaydin et al., 2020).

Despite these challenges, the integration of NNs into stochastic DEA models has shown promise in improving the accuracy and efficiency of efficiency measurement and prediction (Poitier & Cho, 2011). Further research is needed to explore the optimal NN architecture and parameter settings for different types of input and output data, and to investigate the performance of these models on real-world data. Additionally, more research is needed to investigate the interpretability of NN-based stochastic DEA models and how they can be used to support decision-making in practice.

Moreover, some researchers have proposed hybrid models that combine NNs with other machine learning or optimization techniques to further improve the performance of stochastic DEA models. For example, Li et al. (2020) proposed a hybrid model that combines a genetic algorithm with a NN to optimize the input and output weights of the stochastic DEA model, and then uses the optimized weights to calculate the efficiency scores. The results showed that the hybrid model outperformed traditional stochastic DEA models and other machine learning models in terms of prediction accuracy and computational efficiency.

Another challenge of integrating NNs into stochastic DEA models is the issue of interpretability. While NNs are known for their ability to learn complex patterns in data, they are often considered as black-box models because it can be difficult to understand how they arrive at their predictions. This lack of interpretability can be problematic in applications where transparency and accountability are important. However, several methods have been proposed to enhance the interpretability of NN-based stochastic DEA models, such as using feature importance analysis or visualization techniques to identify the most important input and output variables that contribute to the efficiency scores.

Research Problem Statement

The problem of measuring the efficiency of decision-making units (DMUs) under uncertainty is an important challenge in many fields, including finance, healthcare, and manufacturing. While stochastic DEA models have been developed to address this challenge by incorporating uncertainty into efficiency measurement, these models still rely on assumptions about the probability distributions of the input and output variables, which may not always hold true in practice. Moreover, traditional stochastic DEA models may not be able to capture the non-linear relationships between the input and output variables of the DMUs, which can lead to inaccurate efficiency scores.

To overcome these challenges, researchers have proposed integrating neural networks (NNs) into stochastic DEA models to improve the accuracy and efficiency of efficiency measurement and prediction. The basic idea is to use the NN to estimate the distribution of the input and output data of the DMUs, and then use the stochastic DEA model to calculate the efficiency scores based on these estimated distributions.

However, there are several challenges associated with integrating NNs into stochastic DEA models, including selecting an appropriate NN architecture and training algorithm, handling interpretability issues, and dealing with the need for large amounts of training data.

Therefore, the research problem statement is to explore the effectiveness of integrating NNs into stochastic DEA models to improve the accuracy and efficiency of efficiency measurement and prediction under uncertainty. This involves investigating the optimal NN architecture and parameter settings for different types of input and output data, developing methods to enhance the interpretability of NN-based stochastic DEA models, and exploring the performance of these models on real-world data. By addressing these challenges, this research can contribute to the development of more accurate and efficient methods for measuring the efficiency of DMUs under uncertainty, which can have important applications in various fields.

Novelty of Research

The novelty of integrating neural networks (NNs) into the stochastic DEA model lies in the ability to address some of the limitations of traditional DEA models and improve the accuracy and efficiency of efficiency measurement and prediction. By integrating NNs, the stochastic DEA model can capture the complex and non-linear relationships between the input and output variables of the decision-making units (DMUs) and handle uncertainty in the input and output data. This is achieved by using the NN to estimate the distribution of the input and output data and then using the stochastic DEA model to calculate the efficiency scores based on these estimated distributions.

Furthermore, the integration of NNs into the stochastic DEA model allows for the development of hybrid models that combine the strengths of both techniques. For example, some researchers have proposed using genetic algorithms or other optimization techniques to optimize the input and output weights of the stochastic DEA model, which are then used to calculate the efficiency scores based on the estimated distributions from the NN. This results in a more accurate and efficient efficiency measurement and prediction model.

Another novelty of integrating NNs into the stochastic DEA model is the potential for enhancing the interpretability of the model. While NNs are often considered as black-box models, several methods have been proposed to enhance the interpretability of NN-based stochastic DEA models. These methods include using feature importance analysis or visualization techniques to identify the most important input and output variables that contribute to the efficiency scores.

Overall, the integration of NNs into the stochastic DEA model represents a novel approach to addressing the limitations of traditional DEA models and improving the accuracy and efficiency of efficiency measurement and prediction under uncertainty. The development of hybrid models and methods to enhance interpretability further add to the novelty and potential impact of this research.

Plan for the results and discussion of this research

Since there are numerous studies on integrating neural networks into stochastic DEA models, the results and discussions of each study may vary depending on their specific research questions and methodology. However, some common findings and discussions can be highlighted.

One of the main advantages of integrating neural networks into stochastic DEA models is that it can improve the accuracy of efficiency scores. Neural networks can capture the complex relationships between input and output variables more effectively than traditional DEA models, which assume a linear relationship between inputs and outputs. This improved accuracy can lead to better decision-making in various applications, such as selecting the most efficient suppliers, evaluating corporate social responsibility performance, and measuring environmental efficiency.

In addition to improving accuracy, integrating neural networks can also help address some of the limitations of traditional DEA models. For example, traditional DEA models assume that all inputs and outputs are equally important, which may not be true in real-world applications. By incorporating weighting schemes based on neural networks, researchers can assign different weights to inputs and outputs based on their relative importance, which can lead to more accurate efficiency scores.

Another advantage of integrating neural networks into stochastic DEA models is that it can handle uncertainty and variability more effectively. Traditional DEA models assume that all inputs and outputs are known with certainty, which may not be the case in many real-world applications. By incorporating stochastic elements into DEA models and using neural networks to estimate probability distributions of inputs and outputs, researchers can account for uncertainty and variability more effectively.

However, there are also some challenges and limitations to integrating neural networks into stochastic DEA models. One challenge is that neural networks require large amounts of data to train effectively, which may be a limitation in some applications. Additionally, neural networks are often regarded as "black boxes," meaning that it may be difficult to interpret the results and understand the underlying logic of the models.

In conclusion, integrating neural networks into stochastic DEA models has the potential to improve accuracy, address limitations of traditional DEA models, and handle uncertainty and variability more effectively. However, more research is needed to fully explore the advantages and limitations of these models and to develop effective methodologies for specific applications.

Reference

Aggarwal, I., Gunreddy, N., & Rajan, A. J. (2021). A Hybrid Supplier Selection Approach Using Machine Learning and Data Envelopment Analysis. *2021 Innovations in Power and Advanced Computing Technologies (i-PACT)*, 1–5. https://doi.org/10.1109/i-PACT52855.2021.9696826

Amirkhan, M., Didehkhani, H., Khalili-Damghani, K., & Hafezalkotob, A. (2018). Mixed uncertainties in data envelopment analysis: A fuzzy-robust approach. *Expert Systems with Applications*, *103*(1), 218–237. https://doi.org/https://doi.org/10.1016/j.eswa.2018.03.017

Amirteimoori, A., Charles, V., & Mehdizadeh, S. (2023). Stochastic data envelopment analysis in the presence of undesirable outputs. *Journal of the Operational Research Society*, *74*(2), 1–14. https://doi.org/https://doi.org/10.1080/01605682.2023.2172366

Ang, S., Zhu, Y., & Yang, F. (2021). Efficiency evaluation and ranking of supply chains based on stochastic multicriteria acceptability analysis and data envelopment analysis. *International Transactions in Operational Research*, *28*(6), 3190–3219. https://doi.org/https://doi.org/10.1111/itor.12707

Angiz, M. Z., Mustafa, A., & Kamali, M. J. (2013). Cross-ranking of decision making units in data envelopment analysis. *Applied Mathematical Modelling*, *37*(1–2), 398–405. https://doi.org/https://doi.org/10.1016/j.apm.2012.02.038

Apaydin, H., Feizi, H., Sattari, M. T., Colak, M. S., Shamshirband, S., & Chau, K.-W. (2020). Comparative analysis of recurrent neural network architectures for reservoir inflow forecasting. *Water*, *12*(5), 1500. https://doi.org/https://doi.org/10.3390/w12051500

Aydın, Z., & Toklu, B. (2023). Stochastic Data Envelopment Analysis in Measuring the Efficiency of Electricity Distribution Companies. In *Energy Systems Design for Low-Power Computing* (pp. 305–334). IGI Global. https://doi.org/10.4018/978-1-6684-4974-5.ch015

Balak, S., Behzadi, M. H., & Nazari, A. (2021). Stochastic copula-DEA model based on the dependence structure of stochastic variables: An application to twenty bank branches. *Economic Analysis and Policy*, *72*(12), 326–341. https://doi.org/https://doi.org/10.1016/j.eap.2021.09.002

Boubaker, S., Le, T. D. Q., Ngo, T., & Manita, R. (2023). Predicting the performance of MSMEs: a hybrid DEA-machine learning approach. *Annals of Operations Research*, 1–23. https://doi.org/https://doi.org/10.1007/s10479-023-05230-8

Bouzidi, Z., Amad, M., & Boudries, A. (2022). Deep Learning-Based Automated Learning Environment Using Smart Data to Improve Corporate Marketing, Business Strategies, Fraud Detection in Financial Services, and Financial Time Series Forecasting. *International Conference on Managing Business Through Web Analytics*, 353–377. https://doi.org/10.1007/978-3-031-06971-0_26

Brumm, P., Ciotta, N., Sauer, H. M., Blaeser, A., & Dörsam, E. (2023). Deep learning study of induced stochastic pattern formation in the gravure printing fluid splitting process. *Journal of Coatings Technology and Research*, *20*(1), 51–72. https://doi.org/https://doi.org/10.1007/s11998-022-00687-x

Cinaroglu, S. (2023). Fuzzy Efficiency Estimates of the Turkish Health System: A Comparison of Interval, Bias-Corrected, and Fuzzy Data Envelopment Analysis. *International Journal of Fuzzy Systems*, 1–24. https://doi.org/https://doi.org/10.1007/s40815-023-01519-9

Dalei, N. N., & Joshi, J. M. (2023). Operational efficiency assessment of oil refineries using data envelopment analysis and Tobit model: evidence from India. *International Journal of Energy Sector Management*, *17*(3), 437–454. https://doi.org/https://doi.org/10.1108/IJESM-07-2020-0024

Damayanti, Y., Nainggolan, S., Nurchaini, D. S., & Rahmawati, S. E. (2023). Technical Efficiency Analysis of Fertilizer use for Oil Palm Plantations Self-Help Patterns in Muaro Jambi Regency using Methods Data Envelopment Analysis. *International Journal of Horticulture, Agriculture and Food Science (IJHAF)*, *7*(1), 8–14. https://doi.org/https://dx.doi.org/10.22161/ijhaf.7.1.2

Dotoli, M., Epicoco, N., Falagario, M., & Sciancalepore, F. (2016). A stochastic cross-efficiency data envelopment analysis approach for supplier selection under uncertainty. *International Transactions in Operational Research*, *23*(4), 725–748. https://doi.org/https://doi.org/10.1111/itor.12155

Esmaeilzadeh, P. (2020). Use of AI-based tools for healthcare purposes: a survey study from consumers' perspectives. *BMC Medical Informatics and Decision Making*, *20*(1), 1–19. https://doi.org/https://doi.org/10.1186/s12911-020-01191-1

Fallahpour, A., Olugu, E. U., Musa, S. N., Khezrimotlagh, D., & Wong, K. Y. (2016). An integrated model for green supplier selection under fuzzy environment: application of data envelopment analysis and genetic programming approach. *Neural Computing and Applications*, *27*(4), 707–725. https://doi.org/https://doi.org/10.1007/s00521-015-1890-3

Gholizadeh, H., Fathollahi-Fard, A. M., Fazlollahtabar, H., & Charles, V. (2022). Fuzzy data-driven scenario-based robust data envelopment analysis for prediction and optimisation of an electrical discharge machine's parameters. *Expert Systems with Applications*, *193*(5), 116419. https://doi.org/https://doi.org/10.1016/j.eswa.2021.116419

Habib, M. S., Omair, M., Ramzan, M. B., Chaudhary, T. N., Farooq, M., & Sarkar, B. (2022). A robust possibilistic flexible programming approach toward a resilient and cost-efficient biodiesel supply chain network. *Journal of Cleaner Production*, *366*(15), 132752. https://doi.org/https://doi.org/10.1016/j.jclepro.2022.132752

Joro, T., Korhonen, P., & Wallenius, J. (1998). Structural comparison of data envelopment analysis and multiple objective linear programming. *Management Science*, *44*(7), 962–970. https://doi.org/https://doi.org/10.1287/mnsc.44.7.962

Kainthura, P., & Sharma, N. (2022). Hybrid machine learning approach for landslide prediction, Uttarakhand, India. *Scientific Reports*, *12*(1), 20101. https://doi.org/https://doi.org/10.1038/s41598-022-22814-9

Khaki, A. R., Sadjadi, S. J., Gharakhani, M., & Rashidi, S. (2012). Data envelopment analysis under uncertainty: A case study from public healthcare. *African Journal of Business Management*, *6*(24), 7096–7105. https://doi.org/10.5897/AJBM11.591

Labijak-Kowalska, A., & Kadziński, M. (2023). Exact and stochastic methods for robustness analysis in the context of Imprecise Data Envelopment Analysis. *Operational Research*, *23*(1), 22. https://doi.org/https://doi.org/10.1007/s12351-023-00755-z

Maniati, M., Sambracos, E., & Sklavos, S. (2022). A Neural Network approach for integrating banks' decision in shipping finance. *Cogent Economics & Finance*, *10*(1), 2150134. https://doi.org/https://doi.org/10.1080/23322039.2022.2150134

Modhej, D., & Dahimavi, A. (2022). Evaluation Efficiency of Large-Scale Data Set: Cerebellar Model Articulation Controller Neural Network. *Iranian Journal of Operations Research*, *13*(1), 13–30. https://iors.ir/journal/article-1-771-en.pdf

Namakin, A., Najafi, S. E., Fallah, M., & Javadi, M. (2021). A New Hybrid Methodology Based on Data Envelopment Analysis and Neural Network for Optimization of Performance Evaluation. *International Journal of Industrial Mathematics*, *13*(4), 395–409. https://doi.org/http://dorl.net/dor/20.1001.1.20085621.2021.13.3.4.1

Nazari-Shirkouhi, S., Tavakoli, M., Govindan, K., & Mousakhani, S. (2023). A hybrid approach using Z-number DEA model and Artificial Neural Network for Resilient supplier Selection. *Expert Systems with Applications*, *222*(7), 119746. https://doi.org/https://doi.org/10.1016/j.eswa.2023.119746

Pendharkar, P. C. (2023). A Radial Basis Function Neural Network for Stochastic Frontier Analyses of General Multivariate Production and Cost Functions. *Neural Processing Letters*, 1–22. https://doi.org/https://doi.org/10.1007/s11063-022-11137-5

Poitier, K., & Cho, S. (2011). Estimation of true efficient frontier of organisational performance using data envelopment analysis and support vector machine learning. *International Journal of Information and Decision Sciences*, *3*(2), 148–172. https://doi.org/https://doi.org/10.1504/IJIDS.2011.040421

Qu, S., Feng, C., Jiang, S., Wei, J., & Xu, Y. (2022). Data-Driven Robust DEA Models for Measuring Operational Efficiency of Endowment Insurance System of Different Provinces in China. *Sustainability*, *14*(16), 9954. https://doi.org/https://doi.org/10.3390/su14169954

Sengupta, J. K. (1987). Data envelopment analysis for efficiency measurement in the stochastic case. *Computers & Operations Research*, *14*(2), 117–129. https://doi.org/https://doi.org/10.1016/0305-0548(87)90004-9

Shi, Y., & Zhao, W. (2023). An Integrated machine learning and DEA-predefined performance outcome prediction framework with high-dimensional imbalanced data. *INFOR: Information Systems and Operational Research*, *61*(1), 1–30. https://doi.org/https://doi.org/10.1080/03155986.2023.2168943

Tang, Y., Song, Z., Zhu, Y., Yuan, H., Hou, M., Ji, J., Tang, C., & Li, J. (2022). A survey on machine learning models for financial time series forecasting. *Neurocomputing*, *512*(11), 363–380. https://doi.org/https://doi.org/10.1016/j.neucom.2022.09.003

Wang, T., Wang, X., Jiang, Y., Sun, Z., Liang, Y., Hu, X., Li, H., Shi, Y., Xu, J., & Ruan, J. (2022). Hybrid machine learning approach for evapotranspiration estimation of fruit tree in agricultural cyber-physical systems. *IEEE Transactions on Cybernetics*. https://doi.org/10.1109/TCYB.2022.3164542

Wu, J., Shen, L., Zhang, G., Zhou, Z., & Zhu, Q. (2022). Efficiency evaluation with data uncertainty. *Annals of*

Operations Research, *22*(04636), 1–25. https://doi.org/https://doi.org/10.1007/s10479-022-04636-0

Yazdanparast, R., Tavakkoli-Moghaddam, R., Heidari, R., & Aliabadi, L. (2021). A hybrid Z-number data envelopment analysis and neural network for assessment of supply chain resilience: a case study. *Central European Journal of Operations Research*, *29*(12), 611–631. https://doi.org/https://doi.org/10.1007/s10100-018-0596-x

Zarrin, M., & Brunner, J. O. (2023). Analyzing the accuracy of variable returns to scale data envelopment analysis models. *European Journal of Operational Research*, *308*(3), 1286–1301. https://doi.org/https://doi.org/10.1007/s40815-023-01519-9

Zhu, N., Zhu, C., & Emrouznejad, A. (2021). A combined machine learning algorithms and DEA method for measuring and predicting the efficiency of Chinese manufacturing listed companies. *Journal of Management Science and Engineering*, *6*(4), 435–448. https://doi.org/https://doi.org/10.1016/j.jmse.2020.10.001