Data driven approach for stochastic DEA in machine learning and artificial intelligence to improve the accuracy, stability, and interpretability of the model

Hengki Tamando Sihotang¹, Zhimin Huang², Aisyah Alesha²
¹Institute of Computer Science, Indonesia
²School of Management and Business, Adelphi University, USA

Introduction

Stochastic DEA (Data Envelopment Analysis) is a widely used method for measuring the relative efficiency of decision-making units (DMUs) based on multiple inputs and outputs (Saen et al., 2005)(X. Zhang et al., 2023). The incorporation of stochastic elements into the DEA model allows for the measurement of efficiency in the presence of uncertainty and variability in the data (Wanke et al., 2023)(Izadikhah, 2022). However, traditional DEA models may not be able to capture complex and nonlinear relationships between inputs and outputs (Umenweke et al., 2022). Recent advances in machine learning and artificial intelligence (AI) have provided opportunities to improve the accuracy, stability, and interpretability of stochastic DEA models (Heng & Subramanian, 2022).

One approach is to integrate deep learning techniques, such as neural networks, into the stochastic DEA model (Tayal et al., 2020). Neural networks can learn complex and nonlinear relationships between inputs and outputs, which may be missed by traditional DEA models (Q. Zhang et al., 2018)(Dawson & Wilby, 1998). By incorporating neural networks into the stochastic DEA framework, we can improve the prediction accuracy of input-output data, especially for large datasets (Yousefi et al., 2023).

Another approach is to use ensemble learning techniques, such as Random Forest and Gradient Boosting, to improve the stability and generalizability of the stochastic DEA model (Dehaghani et al., 2022). By combining multiple models, ensemble learning can reduce the variance and improve the predictive power of the model (Lin et al., 2022).

Hybrid models that combine stochastic DEA with other machine learning techniques, such as clustering or dimensionality reduction, can potentially improve the interpretability and comprehensibility of the model while maintaining the efficiency of the stochastic DEA framework (Keshteli & Rostamy-Malkhalifeh, 2022).

Dynamic stochastic DEA models that can adapt to changes in the data over time can provide more accurate and actionable insights for decision-making (Yang et al., 2018).
By incorporating time-series data and accounting for changes in the efficiency of DMUs over time, dynamic stochastic DEA models can provide a more accurate representation of the DMUs' efficiency and their driving factors (Luo et al., 2019).

Finally, incorporating explainable AI techniques into the stochastic DEA model can improve its interpretability. Explainable AI can help identify the key factors driving the efficiency of DMUs, making the model more transparent and easier to understand.

In conclusion, the integration of machine learning and AI techniques into stochastic DEA models can improve their accuracy, stability, and interpretability, leading to more informed and effective decision-making.

**Research Problem Statement**

Despite its popularity, traditional DEA models may not be able to capture complex and nonlinear relationships between inputs and outputs. This limitation can lead to inaccurate efficiency scores and ineffective decision-making. Therefore, there is a need to explore new approaches to improve the accuracy, stability, and interpretability of stochastic DEA models.

Recent advances in machine learning and AI have provided opportunities to address this problem. However, there is a lack of research that explores the integration of these techniques into stochastic DEA models. This research gap presents a problem because decision-makers need accurate and interpretable models to make informed decisions.

Therefore, the problem statement for this research is how to develop a data-driven approach that incorporates machine learning and AI techniques into stochastic DEA models to improve their accuracy, stability, and interpretability. This research aims to bridge the gap between traditional DEA models and advanced machine learning and AI techniques to provide decision-makers with more accurate, reliable, and interpretable models for measuring efficiency and identifying the factors driving efficiency. By addressing this problem, the proposed research can have significant implications for decision-making in various fields, including healthcare, finance, and manufacturing.

**Novelty**

The novelty of this research lies in the integration of machine learning and artificial intelligence techniques into stochastic DEA models. While traditional DEA models have been widely used to measure the efficiency of decision-making units, they may not be able to capture complex and nonlinear relationships between inputs and outputs. By integrating advanced machine learning and AI techniques, this research aims to
improve the accuracy, stability, and interpretability of stochastic DEA models, providing decision-makers with more reliable and actionable insights.

Moreover, this research explores several novel approaches, including the integration of deep learning techniques, ensemble learning, dynamic stochastic DEA models, and explainable AI, to improve the performance of stochastic DEA models. These approaches have the potential to enhance the accuracy of efficiency scores, increase the stability of the model, provide more actionable insights, and improve the model's interpretability.

By integrating these approaches into stochastic DEA models, this research aims to provide a comprehensive and effective solution to the problem of measuring the efficiency of decision-making units. This approach has not been explored extensively in the literature, and thus represents a novel and innovative approach to addressing this important research problem.
Reference


