

Expanding capabilities for policy evaluation through deep learning emulation of CGE models

Deep learning (DL) models, which utilise multi-layered artificial neural networks (ANNs), have emerged as a powerful tool for uncovering complex patterns in data. By leveraging large datasets, DL models have achieved state-of-the-art performance in pattern recognition and prediction tasks, outperforming traditional machine learning methods.

The ability of DL models to capture non-linear relationships between high-dimensional input and output vectors makes them an attractive candidate for emulating complex structural models. Despite their potential, DL methods have seen limited application in economy-wide modeling for decision-making, with most studies focusing on agent-based models and farm-level models. Notably, the use of DL as a metamodel for complex optimization problems, such as computable general equilibrium (CGE) models, remains largely unexplored. This gap in research presents an opportunity to investigate the capabilities of DL models in capturing the structure of large simulation models, with potential implications for improving the accuracy and efficiency of economic modeling and decision-making.

The validation of emulators for structural models has numerous benefits. Firstly, it enables non-technical users to interact with complex models through simpler interfaces, facilitating policy evaluation and decision-making. Secondly, it promotes participatory modeling by bridging the technical gap between experts and stakeholders, allowing for more transparent and collaborative decision-making. Thirdly, open-source emulators can reduce costs and increase computational efficiency, enabling faster convergence of results in multi-model frameworks. Lastly, emulators can perform large-scale sensitivity analyses, overcoming limitations of deterministic models, and serve as metamodels for calibration procedures.

This study aims to demonstrate the reliability of Artificial Neural Networks (ANNs) in emulating complex structural models, such as CGE models. We evaluate the performance of ANNs with a high dimensionality of the input and output spaces (100 inputs and 1500 outputs). Furthermore, the hyperparameter¹ space of ANNs is wide, with numerous parameters that can significantly impact model performance, making manual tuning impractical and time-consuming. To address this challenge, we construct a fine-tuning framework which employs multi-objective evolutionary algorithms (MOEA) to identify the most impactful hyperparameters and their optimal values (Figure 1). By doing so, we aim to not only demonstrate the effectiveness of ANNs in emulating CGE models but also provide valuable lessons for future applications involving similar models, including insights into the most critical hyperparameters and their optimal settings. This will enable researchers and practitioners to develop more efficient and accurate emulators.

¹ Hyperparameters in ANNs refer to the parameters that are set before training the model, such as the number of hidden layers, the number of neurons in each layer, the activation functions, the learning rate or the regularization technique. For example, the number of hidden layers and the number of neurons in each layer determine the complexity of the model, while the learning rate controls how quickly the model learns from the training data.

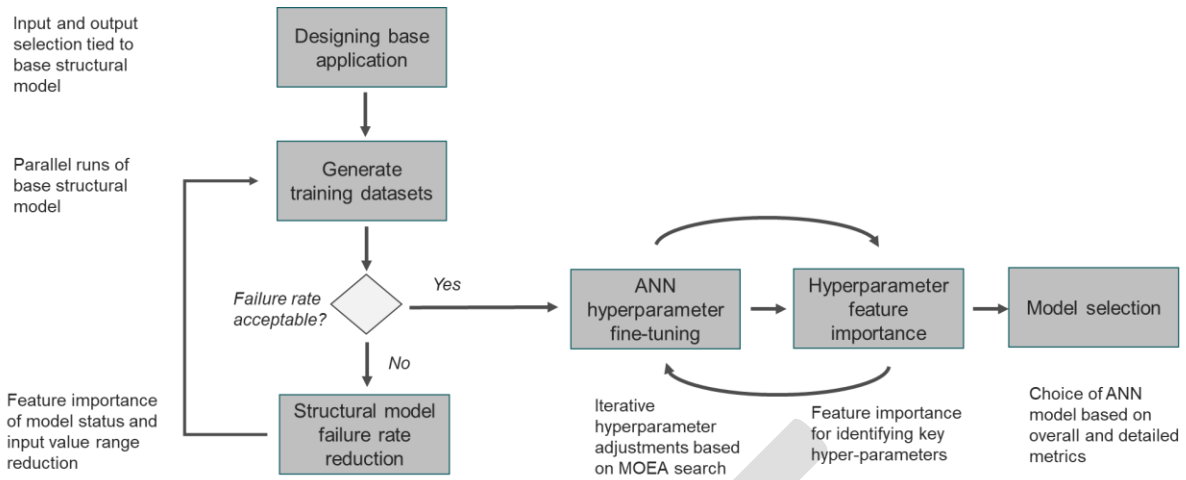


Figure 1 – Overview of the emulator development

The study starts from the DEMETRA model as a base model to be emulated. DEMETRA is a single-country CGE model used to analyse policies for economic development and food security. This model is calibrated with a 2017 Social Accounting Matrix (SAM) for the Kenyan economy, which consists of 71 commodities, 51 activities, 29 factors of production, and 22 household groups. This results in a model size of over 32 thousand equations (with an equal number of free variables). For this application, we aim to emulate DEMETRA to enable policy analysis within the Kenyan context. The possibilities of policy analysis cover various instruments (100 individual model parameters and exogenous variables) that can be combined to analyse different policy scenarios (combining e.g. government spending, fiscal reform, productivity changes). The goal is to develop an extended capability for running policy analysis, allowing for the combination of different instruments within pre-defined ranges. We also employ these ranges to generate the training data for the ANN emulator using Latin Hypercube sampling with uniform distributions. The samples are applied to the CGE model which is ran on a cluster computing facility to generate 100 thousand data points for the training and testing datasets.

In the fine-tuning framework, we define the MOEA problem across three objectives resulting from the ANN training and testing – R^2 on the training set, R^2 on the testing set and the overall inference time. The first two objectives are to be maximized and are considered together to ensure that the model performance is high and that the hyperparameter choice does not lead to overfitting. The third objective which is minimised, is a proxy for model complexity and enables the selection of the simplest model configurations that lead to a high prediction performance. To address the issue of local optima of the MOEA search, we perform searches using different seeds as starting points for the MOEA populations.

Preliminary results

The best performing models determine an R^2 score in excess of 0.995 overall for both test and train sets (Figure 3), highlighting an excellent capacity of multi-layered ANNs to emulate CGE models. The breakdown of this score across the different output nodes, grouped around the structural model variables, indicates some variation, however, still with a lower bound of 0.95. Among the DEMETRA variables, CPI has the lowest performance, while GDP and the exchange rate are more in line with the overall performance score. We also obtain that more hidden layers within the ANN do not necessarily lead to a better performance. The peak R^2 is obtained with two hidden layers, while a model with just one hidden layer still yields similar performance as those with three or more layers.

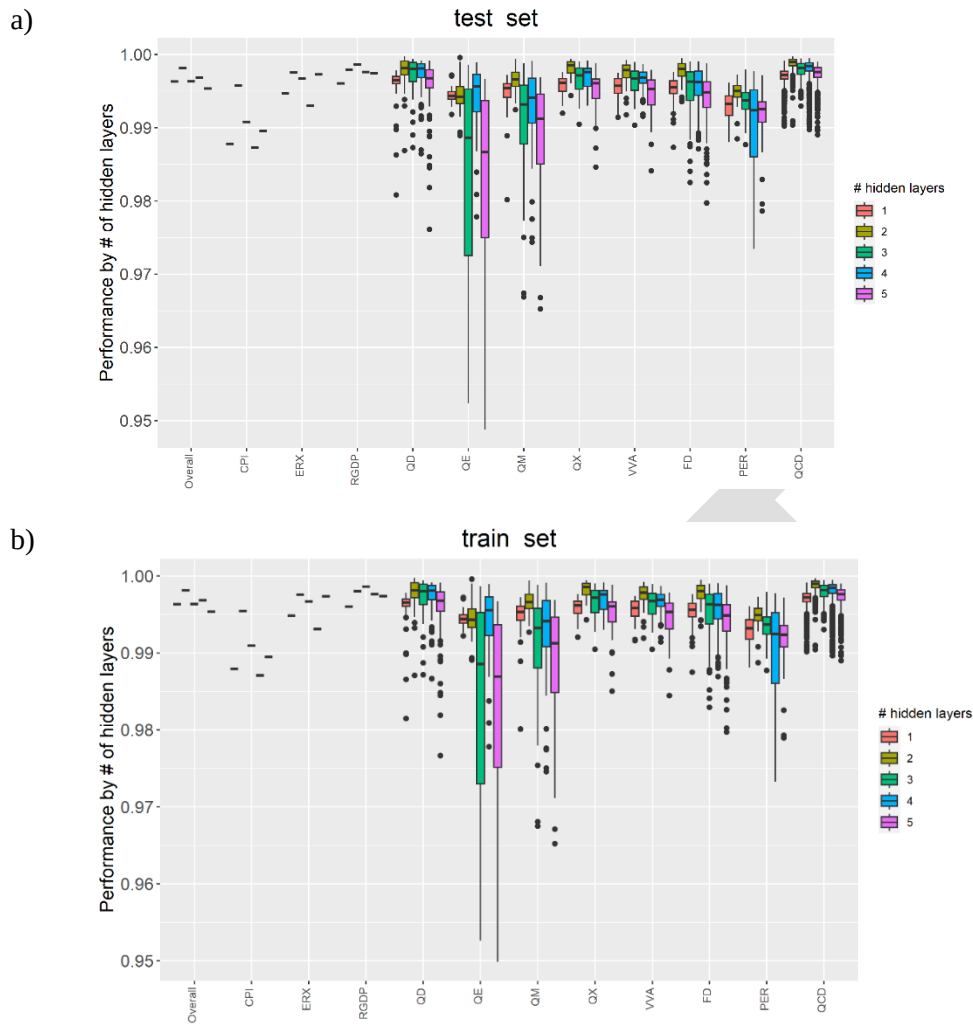


Figure 2 - Top R^2 performance models by number of layers detailed by DEMETRA variables as ANN output nodes. Labels on the x-axis depict different variables of the emulated CGE model.

Next steps

To highlight the computational advantages of emulators, an expansion of this analysis will be to include CGE behavioral parameters (elasticities) as inputs for the ANN. This addition will enable the running of Monte Carlo analyses at a fraction of the time required with more standard approaches. The emulator will also be embedded into an interactive tool allowing non-technical users to engage with the different policy instruments and to retrieve results at different levels of aggregation (Figure 3).

Policy Emulator Kenya

Full screen

Inputs: [Expand](#) / [Collapse](#) / [Reset](#)

Foreign investment

Govt spending

Sales tax scaling

Specific sales tax change

Corp tax scaling

Income tax

Factor use tax change

Export demand shift

Trade margins

Productivity change

Land. Agrifood: 0 (pc)

Capital. Agrifood: 0 (pc)

Labour. Agrifood: 0 (pc)

Capital. Manufacturing: 0.05 (pc)

Labour. Manufacturing: 0 (pc)

Capital. Services: 0 (pc)

Labour. Services: 0 (pc)

Visualizations Tables Dev info

Show output as:

Absolute value

Difference from base

Percentage difference

Select output:

Imports

Macro commodities

All

Imports by Macro commodities

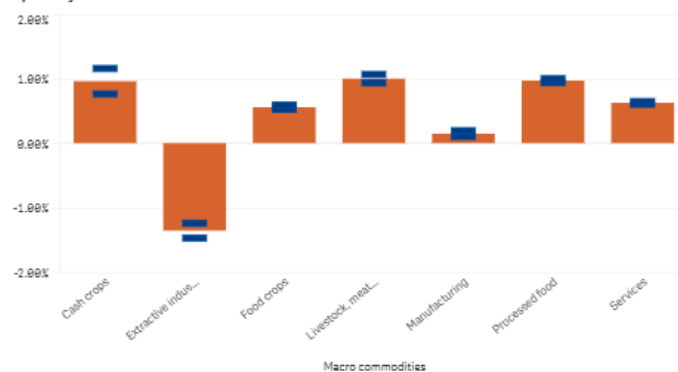


Figure 3 – Screen capture of policy emulator with an example capital productivity shock in manufacturing sectors, embedding a Monte Carlo analysis over trade elasticities