Leveraging Machine Learning in Input-Output Economic Modeling

Topic: Input-Output Theory and Methodology (2)
Author: Kehan He
Co-Authors: Zhifu Mi

The Input-Output (IO) model is instrumental in analyzing economic interdependence by tracking how output from one sector serves as inputs for another. Despite its successful applications, the IO model faces challenges such as time lags and inaccuracies due to the labor-intensive nature of compiling IO tables, which are typically updated only every five years. To address these limitations, this research explores machine learning as a promising alternative for estimating and forecasting economic structures based on the IO model. Machine learning can uncover complex relationships between socio-economic parameters and IO structures without relying on intricate economic theories, offering a cost-effective method for predicting economic changes.

The research assumes that a country's development follows a trajectory defined by socio-economic indicators, such as education level, capital and resource availability, and trade statistics, which in turn determine its economic structure over time. By utilizing publicly available socio-economic data from the annually updated World Bank Development Indicators, the study circumvents the challenges of constructing IO tables through labor-intensive surveys by applying the latest machine learning techniques. Using OECD annual IO tables, the IO tables in this research are aggregated into nine sectorsâ€”Food, Chemical, Mining, Transport, Heavy Industry, Manufacturing, Energy Consumption, Construction, and Serviceâ€”to reduce computational demands. A Random Forest algorithm is employed to predict sector-to-sector production ratios, thereby constructing an IO structure using 829 indicators from the World Bank Development Indicators to train models for each element of a 9x9 IO table for each country.

The results demonstrate the promising potential of this method. When defining relatively accurate predictions as those ranging from 50% to 200% of the true value, 27% of the IO table cell data in the testing set fall within this range. Additionally, the accuracy of predictions varies across different countries and sectors. For example, the service sector shows an accuracy rate of 44%, indicating that its contribution to the economic structure, as described by the IO table, can be more precisely estimated using the World Bank Development Indicators employed in this research.

To further illustrate the model's potential as a viable alternative to complex economic models, a sensitivity analysis was conducted on economic structure changes in response to variations in the renewable energy ratio for Vietnam. The simulation reveals a nonlinear pattern of economic structure change with an increased renewable energy ratio. For instance, when the renewable energy ratio increased from 15% to the current level in Vietnam, the contribution of the chemical sector to energy consumption decreased by 0.4%, while its contribution to itself increased by 0.5%. If the renewable energy ratio were to rise to 30%, the changes in the chemical sector's contributions to all other sectors would be minimal, limited to less than 0.1%.

Overall, this method shows great potential in economic research and represents a significant advancement in IO modeling by integrating it with the latest developments in data science and machine learning. However, the research was constrained by computational hardware limitations, preventing the inclusion of more parameters and a more detailed breakdown of IO sectors. Future advancements could be achieved by incorporating additional parameters from multiple sources and applying better computational devices to enhance the model's performance.