Modelling the Economic Behaviour of a Company Using the Business Accounting Matrix (BAM) and Machine Learning

Topic: Input-Output Theory and Methodology (9)
Author: Casiano A. Manrique de Lara PeÃ±ate
Co-Authors: Javier Lorenzo-Navarro, Jose Juan Deniz Mayor

Modelling the Economic Behaviour of a Company Using the Business Accounting Matrix (BAM) and Machine Learning
Casiano A. Manrique de Lara PeÃ±ate
Jose Javier Lorenzo Navarro
JosÃ© Juan DÃ©niz Mayor

Keywords: Business Accounting Matrices (BAM), Machine learning.

In the current economic context, characterised by high uncertainty and increasingly intense competition, the ability to anticipate the financial behaviour of a company has become a key factor for business survival and growth. Many small and medium-sized enterprises (SMEs) succumb to a market that does not forgive a lack of foresight, while those that manage to adapt and anticipate changes succeed in enduring. In this context, the research question addressed in this work is the following: is it possible to use deep learning techniques, specifically LSTM and CNN-LSTM neural network architectures, to reliably predict the future financial behaviour of Spanish SMEs based on their historical accounting data represented through Business Accounting Matrices (BAM)?

The data used in this study come from the Iberian Balance Sheet Analysis System (SABI) [Informa D&B, n.d.], a web-based tool developed by INFORMA in collaboration with Bureau Van Dijk that provides access to detailed financial information on more than 2.7 million Spanish companies. From SABI, the annual accounts of an initial set of 6,599 Spanish SME hospitality companies were obtained and subsequently transformed into Business Accounting Matrices (BAM) [Manrique-de-Lara-PeÃ±ate & DÃ©niz-Mayor, 2022]. BAMs constitute a matrix-based representation of the economic flows of a company, organised according to categories of economic activity and the agents involved, and are structured into three main blocks: production, distribution, and the use of goods and services. Each cell of the matrix reflects a transaction between departments or accounting entries, providing a comprehensive and concise overview of business activity in a national accounting format. After a thorough cleaning and filtering processâ€”which included the removal of companies with excessive null values, the elimination of outliers with revenues exceeding 300,000 euros, and the exclusion of records with negative sales valuesâ€”the final dataset comprised 1,301 companies and slightly over 10 million data points.

In the modelling phase, two deep neural network architectures were implemented and compared. First, a simple LSTM (Long Short-Term Memory) network [Hochreiter & Schmidhuber, 1997] was employed, chosen for its ability to capture temporal dependencies in sequential data. The second architecture was a CNN-LSTM network, which combines convolutional layers (CNN), capable of extracting spatial features from data, with LSTM layers to model sequential dependencies. The BAM matrices of each company were converted into greyscale images to feed the convolutional component of the model. A hyperparameter search was conducted exploring different combinations of epochs (150, 200, 250), batch sizes (8, 16, 32), and random seeds. The primary evaluation metric was the mean squared error (MSE).

The results obtained show that both architectures are capable of learning relevant patterns from the available financial data. However, the simple LSTM network achieved an MSE of 0.0016, notably

outperforming the CNN-LSTM network, which obtained an MSE of 0.0045. The main novelty of this research lies in the application, for the first time according to the literature review conducted, of LSTM and CNN-LSTM neural network architectures using BAM matrices as input data. While previous studies have employed neural networks for stock price prediction or corporate bankruptcy detection based on conventional financial indicators, this study introduces a structured, matrix-based representation of business accountingâ€”the BAM, proposed by Manrique-de-Lara-PeÃ±ate & DÃ©niz-Mayor (2022)â€”as input for deep learning models. A first attempt in this direction by the authors was carried out in Herrera CÃ¡rdenes et al. (2023), where an initial proof of concept was developed that validated the viability of the approach and laid the methodological foundations for the present work.

This combination enables the capture of not only the temporal evolution of financial variables, but also the interrelationships between the different areas and economic flows of the company. The results, although preliminary given the size of the dataset, constitute an encouraging proof of concept that demonstrates the usefulness of BAM as a rich, structured data source for machine learning applications. More broadly, this work opens a promising window for the convergence of economic accounting of firms and sophisticated prediction models, bridging two disciplines that have traditionally evolved separately IO theory and ML.