Bias-Free Text Evaluations in Micro and SME Credit Scoring using Deep Learning

One critical issue in the development of small companies is providing healthy access to funding. Poor data availability has resulted in slow and cumbersome evaluations that rely heavily on written business plans and the written opinions of credit risk analysts. These evaluations carry the internal bias of the evaluator itself, requiring additional validation steps that increase processing times and the cost of the loan.

This paper presents a method for reducing this cost, based on the use of a Deep Neural Network that quantifies the relationship between the text evaluation and the credit risk of the borrower, removing the internal bias that is intrinsic to text. The model takes as input the TF-IDF transform of the text evaluation, information from the borrower, and information from the evaluator, to give a unique numerical value as output summarizing the evaluation which is bias free across two dimensions: the internal evaluator bias across their own evaluations, and the bias compared to the rest of the evaluator staff.

This bias-free result can then be used in a normal credit scoring setting as a new variable. We benchmark this process against competing credit scoring models: without using text-based information, with text-based information using TF-IDF, and against a two-step evaluation incorporating a second risk analyst validation. The results show the usefulness of our model in the context of modern micro and SME risk evaluation, and the potential of emerging technologies for their use in credit scoring.