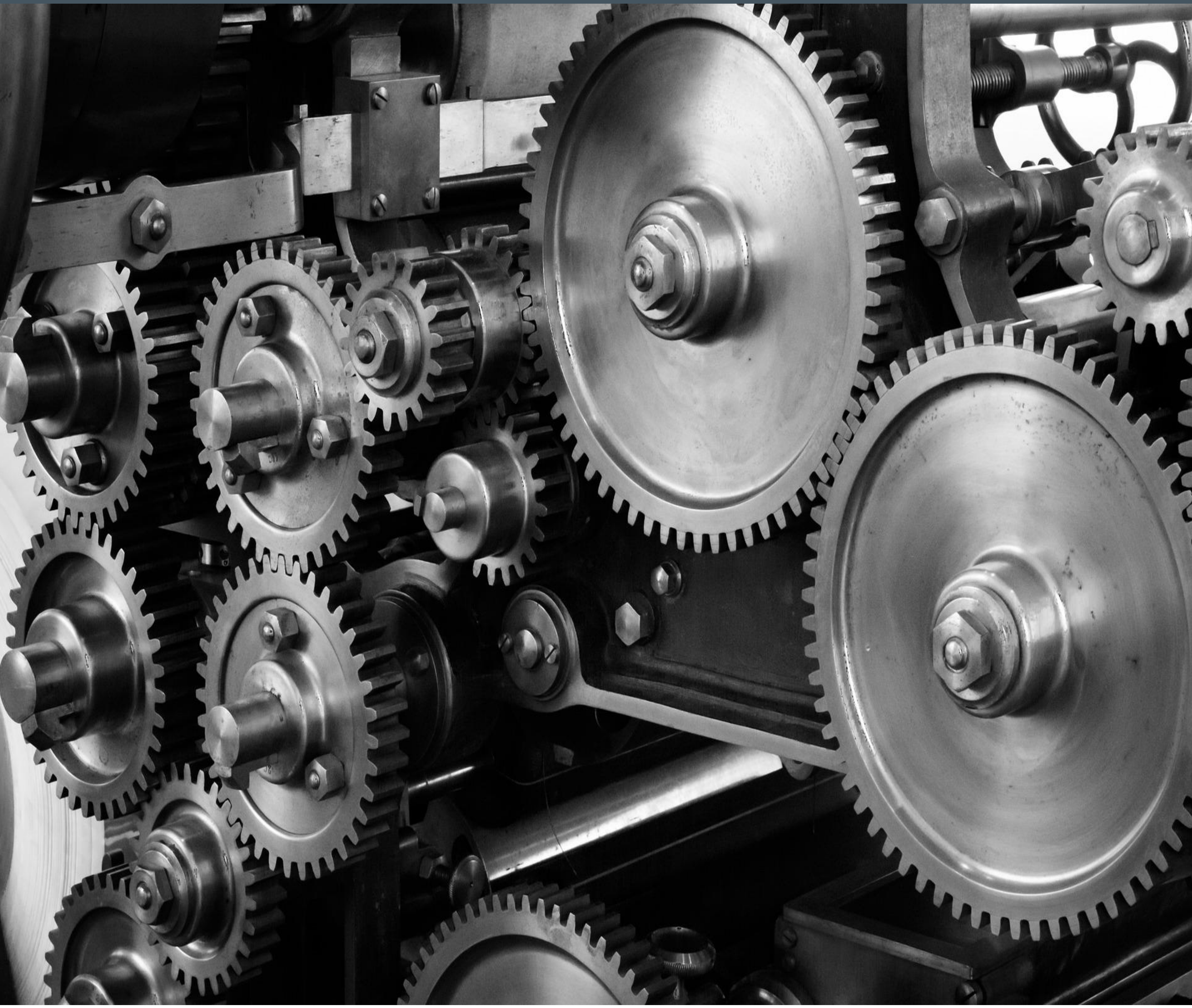
Vol. 1, No. 6 November 2018

New Machinist Journal

Where Machine Learning and Traditional Corporate Finance Meet

Dr. Jimmie Lenz



Financial Risk Group | 264 W. Chatham Street Cary, NC 27511 | 919 439 3819

Machinist (noun): A person who operates a machine, especially a machine tool

The use of a number of different types of machine learning techniques have been discussed at some length in this journal, both to illustrate the breadth of tools in this toolbox, but also the creative results that can be engineered in the right hands. A questioned was recently posed about the application in work that we do every day at the Financial Risk Group, and in particular if it plays into some of the research that is done for new products and/or new problems.

While much of what is seen in the popular press around the use of Machine Learning involves theoretical problems and experimentation, which must be done to advance the knowledge and skillsets of practitioners, there is a significant group of work that is also conducted. A recent opportunity to utilize Machine Learning was brought about by a significant change that has been taking place in asset management, the increasing use of Private Capital Investments.

Private Capital investments are often characterized in more detailed terms: Private Equity, Private Debt, and/or Real Assets, and in some aspects differ significantly from the publicly traded versions of these security types. While these investments have been used since the 1980s, their prevalence has exploded in the last decade resulting in changes to the way assets are allocated for many investment managers. However, as the name implies information on the firms that underlie these assets is not public, and this is problematic on a number of fronts for those interested in investments in this asset class.

There are major hurdles in assessing these investments, including a lack of public information, a reasonably short history, and a limited number of these types of securities, even if the data were publicly available. In the case of investment managers' needs when assessing these investments is that of the cash flows. While this is important for any type of investment, Private Capital investments have peculiarities that make this more important. These investments often have very sporadic cashflows that need to be well understood in order to maintain asset allocations and manage cash balances.

The problem is detailed in a recent paper "Macroeconomic Effects on the Modeling of Private Capital Cash Flows" including the need to find the basis on which cash flows could be forecast with respect to timing and severity. While models did exist, they were dated and reflected little of the current environment, requiring a new approach and analysis. As in most modeling problems in which forecasting is the goal independent variable selection is of the utmost importance, as is the notion of excluded variable bias. This is where Machine Learning played an indispensable part in the success of this project.

The nature of investment assets with limited research conducted on them provides few clues from which significant variables of interest might be selected. In the case of Private Capital investments, with different types of underlying assets, the field of economic and capital market data that could potentially be of interest was quite large. Fortunately access to many of the potentially valuable variables was provided via the Federal Reserve's extensive data cache, which provided a broad spectrum of variables in their data from which to set about selecting those with statistically significant predictive power. Selection of the variables, all four models created are multivariate, was conducted utilizing the techniques provided below.

The variable selection process is both automated and manual. Statistical modeling is an art and a science where expert knowledge and mathematical processes combine to describe the world. Expert knowledge selects a set of independent variables which could understandably have explanatory power on the dependent variable. A subset of variables are preselected utilizing the correlation of the independent and dependent variables of interest. Then, using an out of sample Root Mean Squared Error (RMSE) statistic with the associated measurements of significance, final variables are selected.

In this case specifically a neural network is employed, and although a neural network isn't an algorithm, it provides the structure by which the available data can be analyzed utilizing internal machine learning algorithms (e.g. regression models). The model is trained to achieve the desired outcome—minimizing the error term—using an iterative process. The weights of the internal models are updated during the training optimization. The training process continues until optimal model weightings are found. Once an optimal model has been found, out of sample validation observations are utilized to determine the level of improvement (if any exists) and to provide a "score" related to the improvement observed.

The results, as illustrated in the paper, provide an excellent perspective on the actual application of the technology on a real-world problem. The application of the findings has resulted in a significant improvement on a model that no longer meets the need of its users and provides dynamic characteristics, via the variables identified.