

One of These Things is Not Like the Others

Excerpted from the *Alternative Investment Analyst Review*, Volume 8, Issue 1

The Alternative Investment Analyst Review is an official publication of the CAIA Association. Access to the most current issue is an exclusive benefit of CAIA Membership while archived issues are available to the public in the Perspectives section at CAIA.org.

[The full article may be accessed here.](#)

Central Issue of the Paper

Two hedge funds reporting to follow the same strategy, may have very different return generating processes. As a result, statistical properties of their returns may be quite different. It is beneficial to investors to identify funds that perform differently than others labeled as following same strategy. The difference could provide alpha or indicate trouble that needs further investigation. The Editor's Letter "Machine Learning and Hedge Fund Classification using a Self-Organizing Map" is concerned with the ability of a machine learning tool to make such identifications. It illustrates the ability of self-organizing maps (SOMs), which rely on artificial neural networks (ANNs), to group a set of long-short hedge funds into homogeneous groups.

Approach Employed by Paper

Hossein Kazemi and Satyabrota Das explain in simple terms what ANNs are and how SOMs use them to classify data. To illustrate the use of self-organizing maps, they collect and clean a sample of monthly long-short hedge fund returns and divide them into two time periods: January 2012 to December 2015 and January 2016 to December 2017.

After cleaning and removing return series with missing data, they have 195 funds and calculate various statistics for fund, which can be considered features in the machine learning context: average annualized return, average annualized standard deviation, average maximum drawdown, average correlation with the S&P 500 index and average pairwise correlations. It is easy to visualize groups of similar funds when plotted in a two-dimensional space such as mean returns against correlations with the S&P 500. It is much harder to group the funds accounting for four, five or more features at once unless one uses a self-organizing map. The self-organizing map divides the 195-hedge fund returns into 9 groups for the first time period, and into 9 groups for the second time period. A return correlation analysis investigates the similarity of funds within groups compared to correlation of funds with other groups. A transition matrix is then created that illustrates how funds transitioned from one group to another between the two time periods.

Findings of the Paper

Dominant groups of hedge funds classified by the self-organizing maps had a relatively large number of funds. In the first period the dominant groups were group numbers 7, 8 and 9 with 29, 69, and 71 funds, respectively. The other groups (1 through 6) contain only 1 to 6 funds, except group 2 which has 11 funds. A table of the cross-sectional statistics for each group illustrates that fund groups have different characteristics with respect to these features. In addition, the within group correlations are generally high relative to the correlations of funds with other groups. In the second period most of the funds in one large group moved to another showing they continued to follow the same strategy. Yet another large group from period one was split into four groups in the second period indicating that they were no longer homogenous. In summary, the SOM is very good at - like in the Sesame Street song - determining "if one of these [funds] does not belong."
