Objective: Identify future stock market winners (and losers) applying machine learning techniques in anomaly detection using features suggested in “The Anatomy of A Stock Market Winner” by Marc Reinganum (Financial Analysts Journal, March-April 1988, p16).

Steps:

1. Read the paper “The Anatomy of a Stock Market Winner”.
2. Identify a universe of stocks. Downloaded holdings of the iShares Russell 3000 ETF as a source of companies. This provides companies and tickers. There is the flaw of survivorship bias in that this set of companies misses those that existed in the past but went out of business or merged.
3. Download historical returns from Quandl
4. Download historical fundamental ratios from Quandl
5. Examine data for errors
   1. It’s critical to identify erroneous returns, particularly outliers, because these might be confused with true anomalies. This step may also give insight into the problem itself.
   2. Ideas for identifying errors in the returns.
      1. Absolute return across all data: Sort all returns from high to low. Examine both tails. If errors are found keep looking until we find no more errors. Keep an eye out for reversals (high returns followed by low and vice-versa) as this indicates a bad price in the middle of the data.
      2. Returns by week. Calculate the mean and std
6. Merge historical data
7. Do we have enough data.
8. Define anomalies
   1. +100% in 12 months (or less) / -50%?
   2. Top x%?
   3. What would make a good threshold?

Last saved: