Predicting Tail Events for Equity Indices using Machine Learning Algorithms

Henrik Stutz

Tail-risk is an omnipresent risk in the financial world that can significantly affect the value of a financial position. Especially the pre-market trading hours can be a very challenging environment to find the correct price for a security. Moreover, finding this fair price usually must happen quickly.

This thesis attempted to predict tail events during the closing hours of the Euro Stoxx 50 future index in order to improve prediction for its gap return. The S&P 500, Nikkei 225 and the Hang Seng indexes were used as independent variables. The four minute-by-minute time-series used covered the period between 01.11.1999 and 15.04.2016. The tail events were predicted by means of four machine-learning algorithms: CART, Random Forests, Bagging and Boosting. The gap return was predicted with OLS regressions, and return dependencies were analysed with minimum spanning trees. The results show that extending the OLS regressions with tail event prediction, based on the Random Forest algorithm, outperformed OLS regressions without tail information.