Developing early warning systems for financial crises has been a topic of considerable interest, particularly since the 2008-2009 financial crisis. Detecting financial crises is important not only for financial market participants but also for policy makers and regulators and hence there is broad interest in such systems. Developing such a system is very challenging, as predicting changes, particularly extreme changes, in financial markets is very difficult. One method for developing an early warning system is to develop a stress indicator which measures the extent of stress in a financial market, based on leading indicators. An ideal stress indicator should be able to both explain the incidence of a crisis based on the identified leading indicators as well as predict the onset of the crisis (Frankel and Saravelos 2012). Such indicators may broadly be divided into three categories. The first is the regression method indicators which tests the statistical significance of different leading indicators in determining the incidence or probability of a financial crisis (Eichengreen, Rose and Wypslosz 1995, Frankel and Rose 1996). The second is the signalling approach in which exceedance of threshold values of predetermined indicators is the crisis indicator (Kaminsky, Lizondo and Reinhart 1996). The third and most recent method is the data-driven approach examplified in tools and techniques from machine learning and artificial intelligence, which either determine leading indicator crisis thresholds (Frankel and Wei 2004, Bussiere and Fratszcher 2006) or select the most appropriate indicators (Lin et al 2008, Cevim et al 2014).

A financial crisis may be viewed as an abrupt shift in the state of a financial market as a complex dynamical system. Detecting such shifts in financial markets is particularly difficult due to the inherent noisiness of the data as well as potential non-stationarity, which traditional methods of statistical analysis are not always well placed to handle. Novel approaches in data science such as unsupervised feature learning and deep learning (Langkvist, Karlson and Lutfi 2014) have recently been applied in the time series context in order to reduce dimensionality and extract relevant information. One such approach is Topological Data Analysis, which refers to a combination of statistical, computational, and topological methods that finds shape-like persistent structures in data. Topological Data Analysis has been applied in the context of time series and systems analysis (Gholizadeh and Zadrozny 2018) and in particular in the context of financial market crashes (Gidea and Katz 2018). A key property of Topological Data Analysis is its ability to extract stable features from noisy data and it provides a new econometric tool which appears to complement existing statistical techniques in financial crash detection. A key insight from Gidea and Katz (2018) is the notion that topological data analysis is able to detect the “landscape” of crashes in that the shapes of the financial time series depend on the state of the market whose structural changes topological data analysis is able to capture.

In this paper we develop an early warning system for financial crash detection by incorporating topological data analysis into a machine learning framework. This allows us to build a numerical financial stress indicator and also to “unlock” features, which conventional models may supress due to either data or human biases, help detect a coming financial crisis.

DESCRIPTION OF THE METHODOLOGY

We choose the decision-tree-based XGBoost model as the specific machine learning framework to incorporate TDA into the supervised learning process. The speed, simplicity and transparency of the XGBoost model stands out to make it one of the most suitable frameworks for our purpose. Simply speaking, the modelling exercise is divided into two phases: training phase and testing phase. During the training phase, the model processes considerable amount of data which are, in our case, time series of different descriptors of the market in a given time window) and make probabilistic predictions on how likely market will crash in the near future in each round. Then these predictions are compared to the corresponding ground truth to provide feedback, enabling the model to “learn” and improve its predations until the training session is complete. During the testing phase, the model is no longer “learning” but purely making predictions and these predictions are recorded for us to evaluate the performance of the model after all the testing data are processed and predictions are all made.

In order to effectively evaluate the contribution of TDA

DESCRIPTION OF RESULTS

Over the August to December 2008 period the early warning indicators, which is the aggregate probability of a market crash in the next two weeks, look very different with and without TDA features. The indicator without TDA features peaks at the end of November 2008 well after the major decline in the S&P and is consistently elevated only from the beginning of November. When TDA features are incorporated the indicator shows a major spike in mid-September about two weeks before a 25% decline in the S&P and continues to remain elevated in the month of October between two and four weeks before a 17% decline in November. The TDA based indicator again shows a spike at the end of November together with the non TDA based indicator, which does not presage a major decline, However the TDA based indicator drops much more sharply than the non-TDA based indicator. The market decline in 1998 also provides a good test case and here a similar pattern in observed. The S&P peaked in the middle of July 1998, and the TDA based crisis indicator correctly forecasts the drop with 3 consecutive spikes of crisis probability before the market declined around 20%. The crisis indicator without TDA peaked near the end of the decline and remained at low levels prior to the decline.

NEED A SIMILAR GRAPH FOR 1998 HERE

Overall these finding indicate that the TDA based indicator seems to be both a good early warning indicator of crashes and is also less susceptible to false positives than the indicator without TDA features. We find that this finding is robust over longer time periods for the S&P as well as for the FTSE as well as the Hang Seng and that the false positive rates are comparable to those of machine learning based financial crisis indicators (Chatzis et al 2018).

Another important aspect of the TDA based system is feature selection. There are two aspects to this, the first of which is identification of new features and the second is feature importance. Over both the 1998 and 2008 crash periods the crisis indicator incorporating TDA identifies additional features, over and above those initially identified, which are significant for crash prediction. This suggests that the TDA methodology “unlocks” new features and signals around crashes providing further support for the “landscape” of crashes viewpoint proposed in Gidea and Katz (2018). Feature importance with and without these TDA based features is also completely different. When the TDA based features are omitted the machine learning algorithm weights two or three features, mostly involving short term asset returns, very highly and virtually ignoring most of the others. Once TDA features are added feature weighting is far more uniform and, in particular, longer term returns and term spreads as well as the TDA features all receive higher weightings. Thus the presence of TDA features seems to unlock greater value from existing factors which also indicates that incorporating TDA into a machine learning framework could be beneficial in a number of financial market settings.

FALSE NEGATIVES

The issue of false negatives is also very relevant in the context of crash prediction. While it is important to avoid false positives it is perhaps more important to try to avoid false negatives, as the cost of a false negative which is a crash that was not predicted is likely to be much higher than that of a false positive. The precision recall measure is a standard measure in machine learning which measures false positives against false negatives. The precision recall measure of the model without TDA features for high cut-offs, where the probability of a crash is higher the performance, is much lower than for low cut-offs indicating that the model has a much higher incidence of false negatives indicating that the model is much less able to detect crashes when they are more likely to happen. When TDA features are added the precision recall measure declines far more slowly as a function of cut-off and in the case of the FTSE 100 exhibits an increase in a certain region. Thus adding TDA features seems to smooth the performance of the model as the precision recall measure for the model without TDA is often higher than that with TDA for low levels of the cut-off. Taken together with the earlier discussion of feature weighting, it seems that incorporating TDA in this machine learning framework seems to ameliorate the problem of overfitting which seems particularly important in the context of being able to predict crashes that actually happen. This is an advantage of TDA that has been noted in other contexts, namely the ability to exploit weak signals that are nonetheless persistent.

Taking a longer term perspective over the 25 year period from 1993-2018 for the S&P the highest levels for the TDA based crisis indicator were over the 1995-1999 period followed by 2008. This indicates that the late 1990s saw the highest level of stock market uncertainty as proxied by financial market indicators. The relative weight of TDA indicators was lowest over this period with the one week return for the S&P and gold receiving the highest weights. The TDA based feature weighting increased sharply over the 2004-2009 period, declined over the 2010-2012 period and then increased again to its highest overall level while the relative importance of weekly return for the S&P and particularly gold has been much lower since about 1999. These results seem to suggest a pattern of changing linkages within financial markets with TDA able to detect changes in the “shape” of the market.