Exploiting information network model based for financial forecasting

Tomaso Aste, Paola Cerchiello, Giancarlo Nicola

Abstract In this contribution we aim at combining graphical models and recurrent neural networks for predicting bank stock returns. We propose to leverage the information extracted from graphical models fitted on financial data timeseries as input features to a neural network to improve the predictive performance. We focus on the 74 largest listed U.S. banks over a period that spans from 2003 to 2017. We apply a recent and fast algorithm (LoGo) for the network estimation that allows us to calculate the banks' partial correlations network and relative indexes like mutual information and transfer entropy. The information obtained from the network are then leveraged as features in a recurrent neural network and in other predictive models for comparison.

Abstract *In questo lavoro miriamo a combinare modelli grafici e neural network per prevedere i rendimenti delle azioni bancarie. Proponiamo l'integrazione delle informazioni estratte da modelli grafici applicati a serie storiche finanziare come input di una rete neurale per migliorarne le performance predittive. Ci concentriamo sulle 74 maggiori banche quotate statunitensi su un periodo dal 2003 al 2017. Per stimare il modello grafico applichiamo un algoritmo presentato di recente (LoGo) ottenendo il network delle correlazioni parziali e diverse grnadezze come la mutual information e la transfer entropy. Le informazioni ottenute dal modello grafico vengono quindi impiegate come input in una rete neurale ricorrente e in altri modelli predittivi per il realtivo confronto.*

Key words: network model, mutual information, forecast, LoGo algorithm

---

Tomaso Aste
University College London e-mail: t.aste@ucl.ac.uk

Paola Cerchiello
University of Pavia, via San Felice 5 Pavia, e-mail: paola.cerchiello@unipv.it

Giancarlo Nicola
University of Pavia, via San Felice 5 Pavia e-mail: giancarlo.nicola01@universitadipavia.it

1 Introduction

It is well known that market prices are formed in complex interaction mechanisms that, often reflect speculative behaviours rather than the fundamentals of the companies. Market models and, specifically, financial network models based on market data may, therefore, reflect 'spurious' components that could bias systemic risk estimation. This weakness of the market suggests to enrich financial market data with data coming from other, complementary, sources. Indeed, market prices are only one of the evaluations that are carried out on financial institutions: other relevant ones include ratings issued by rating agencies, reports of qualified financial analysts, and opinions of influential media. Insofar, we propose to combine financial data with a sentiment index produced by Reuters, whose time series are integrated by means of a graphical network model and then properly used to predict price time-series.

2 Graphical network model

Here, we briefly describe the graphical network models that will be used to estimate relationships between the N banks, both with market and sentiment data. Direct relationships between banks can be measured by their partial correlation, that expresses the direct influence of a bank on another. Partial correlations can be estimated assuming that the observations follow a graphical Gaussian model, in which the covariance matrix Σ is constrained by the conditional independences described by a graph (see e.g. Lauritzen, 1996). More formally, let $X = (X_1, ..., X_N) \in R^N$ be a $N-$dimensional random vector distributed according to a multivariate normal distribution $\mathcal{N}(\mu, \Sigma)$. Without loss of generality, we will assume that data are generated by a stationary process, and, therefore, $\mu = 0$. In addition, we will assume throughout that the covariance matrix Σ is not singular.

Let $G = (V, E)$ be an undirected graph, with vertex set $V = \{1, ..., N\}$, and edge set $E = V \times V$, a binary matrix, with elements e_{ij}, that describe whether pairs of vertices are (symmetrically) linked between each other ($e_{ij} = 1$), or not ($e_{ij} = 0$). If the vertices V of this graph are put in correspondence with the random variables $(X_1, ..., X_N)$, the edge set E induces conditional independence on X via the so-called Markov properties (see e.g. Lauritzen, 1996). In particular, the pairwise Markov property determined by G states that, for all $1 \leq i < j \leq N$:

$$e_{ij} = 0 \Longleftrightarrow X_i \perp X_j | X_{V \backslash \{i,j\}}; \tag{1}$$

that is, the absence of an edge between vertices i and j is equivalent to independence between the random variables X_i and X_j, conditionally on all other variables $x_{V \backslash \{i,j\}}$.

Let the elements of Σ^{-1}, the inverse of the covariance matrix, be indicated as $\{\sigma^{ij}\}$. Whittaker (1990) proved that the following equivalence also holds:

$$X_i \perp X_j | X_{V \backslash \{i,j\}} \Longleftrightarrow \rho_{ijV} = 0 \tag{2}$$

where

$$\rho_{ijV} = \frac{-\sigma^{ij}}{\sqrt{\sigma^{ii}\sigma^{jj}}} \tag{3}$$

denotes the ij-th partial correlation, that is, the correlation between X_i and X_j, conditionally on the remaining variables $X_{V\setminus\{i,j\}}$.

3 LoGo algorithm

Since we cope with a relatively large number of banks (74), we take advantage of a recently presented algorithm LoGo (Barfuss et al. 2016) to estimate graphical models on the basis of time series data. LoGo is a methodology that makes use of information filtering networks to produce probabilistic models that are sparse, with high likelihood and computationally fast, making possible their usage with very large data sets. LoGo produces high-dimensional sparse inverse covariances by using low-dimensional local inversions, making the procedure computationally efficient and only slightly sensitive to the curse of dimensionality. The construction of the algorithm, through a sum of local inversion, makes this method particularly suitable for parallel computing and dynamical adaptation by local, partial updating, as described in (Barfuss et al. 2016) where a more detailed explanation of the method is presented.

4 Recurrent Neural Network

To forecast the next day (t+1) stock return given the previous day data, we apply a Recurrent Neural Network (RNN). A RNN is a neural network architecture suitable for dynamic problems with temporal-like behaviour like sequences and timeseries. In particular we use a RNN with Long Short-Term Memory cells (LSTM) introduced in (Hochreiter and Schmidhuber 1997). The RNN receives in input the bank stock return timeseries along with the timeseries of the features extracted from the graphical model (e.g. pagerank, mutual information, transfer entropy). The output of the network can be both a classification or a regression depending on the network's output layer definition and on the applied loss-function (Categorical Cross-Entropy for classification and Mean Square Error (MSE) for regression). The network is trained with the Adam optimizer (Kingma and Ba 2014) a variation of the stochastic gradient descent algorithm.

5 Data

For sake of comparability and homogeneity, we focus on a single banking market, the U.S. banking system, characterised by many important banks which have seriously impacted the country economy during the financial crisis. We take into account large listed banks, for which there exist daily financial market data to be compared and integrated with sentiment data produced by Reuters, a well-known news provider.

For each bank, we consider the daily return obtained from the closing price of financial markets, for a period of 3716 days from Jan 2003 through October 2017, as follows:

$$R_t = log(P_t/P_{t-1}) \tag{4}$$

where t is a day, $t-1$ the day preceding it and P_t P_{t-1} the corresponding closing price of that bank in that day.

The sentiment data on the banks, calculated by Reuters, cover the period December 2010 through September 2017. The different time coverage compared to stock data is due to the limited availability of sentiment monitoring services before 2010. The sentiment index is represented by continuous values ranging from -4 to 3 with higher values corresponding to more positive sentiments. For each bank we have then calculated a sentiment daily variation that mimics market returns, as follows:

$$S_t = log(T_t/T_{t-1}) \tag{5}$$

where t is a day, $t-1$ the day preceding it, and T_t T_{t-1} is the corresponding average daily sentiment on that bank for that day.

6 Results

Initially, we calculated a network of the U.S. banks over the entire time horizon to have insights regarding the most correlated banks (from the stock price point of view). Interestingly from figure 1 we can see that the estimated network posits many of the largest bank like C, BAC, GS, MS, TD close to each other and connected by edges in the lower left corner of the network. At the same time, foreign banks like BBVA, BCS, DB, UBS are located together in the right top corner.

Moreover, the LoGo algorithm speed and scalability allows to calculate a different graphical model for each market day based on the data of the previous 90 days. Thus, for every market day we end up with a network that is representative of the bank stocks correlation and market structure in the previous 90 days. This will be important on a later stage to leverage recent market structure information for the predictive task and to develop a dynamic 'snapshot' of how the U.S. bank system relations evolve during the considered period. In fact, it is possible to extract several bank related measures from the graphical model (namely mutual information,

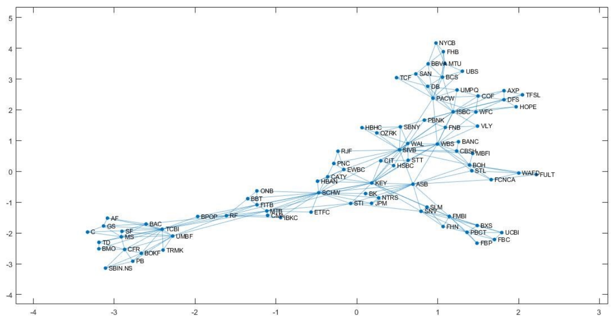

Fig. 1 Resulting graphical network model based on financial data

pagerank, transfer entropy from lagged variables) that allow for an interesting inspection of the system evolution. For example, from figure 2 it is possible to see how the network total mutual information resembles very closely (especially in the trends) the Saint Luis Fed Financial Stress Index (STLFSI). The STLFSI measures the degree of financial stress in the markets and is constructed on 18 weekly data series: seven interest rate series, six yield spreads and five other indicators.

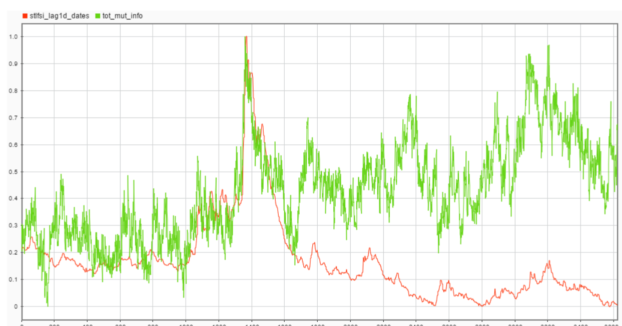

Fig. 2 Trends comparison

We have tested the models on two different prediction tasks to compare their performance. The first task consists in forecasting the return value of the next day closure price given the data from the previous n days (stock return regression). The second task instead entails classifying whether the next day closure price return will be positive or not given the data from the n previous days (stock return classification). Preliminary results on both the tasks show that the combined use of stock returns data and sentiment data holds predictive power for the following day closure price return. We report the results for the model using both stock returns and sentiment data from the 10 previous days on the stock of City bank (ticker C). The model achieves a MSE of 4,40E-4 on the first task (stock return regression) with results on the test set reported in figure 3 (left) and a classification accuracy of 57,1% corresponding to an AUC ROC of 61% (see figure 3 right) for the second task (stock return classification). In the second task the class distribution of the test set is 50% positive and 50% negative returns, thus an accuracy of 57,1% is an improvement compared to a random guess.

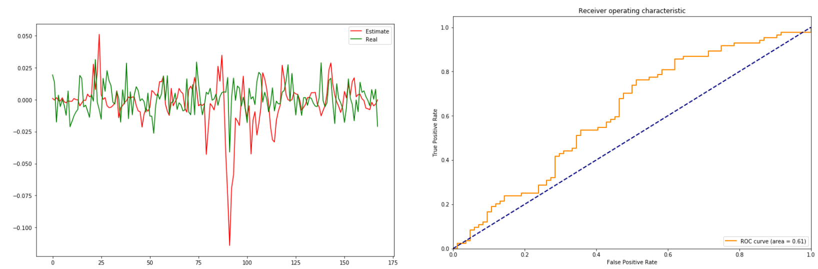

Fig. 3 Prediction results for stock prices (left), Roc curve for stock return classification (right)

7 Conclusions

In this paper we have presented two main contributions. Firstly, we have applied a recently presented graphical model inference methodology to the investigation of U.S. Banks stock returns and sentiment time series to understand the network structure and evolution. Secondly, we have presented a way to leverage the graphical models information for a predictive task on the stock returns with a recurrent neural network and compared the results to other models. The market structure inferred by the graphical models moreover has shown to be an interesting tool for monitoring the U.S. bank system evolution due to its correlation with several known stress measures like the Saint Louis Fed Financial Stress Index.

References

1. Barfuss Wolfram, Massara G. Previde, Di Matteo T. and Aste Tomaso (2016) Parsimonious modeling with information filtering networks. Physical Review E, 94 (6). ISSN 1539-3755
2. Lauritzen, S. L. (1996) Graphical models. Oxford University Press.
3. Whittaker J. (1990) Graphical models in applied multivariate statistics. Wiley Publishing.
4. Sepp Hochreiter and Jurgen Schmidhuber (1997) Long Short-Term Memory. Neural Computation, 9(8):17351780. doi: 10.1162/neco.1997.9.8.1735.
5. Kingma, Diederik P. and Ba, Jimmy (2014) Adam: A Method for Stochastic Optimization. arXiv:1412.6980 [cs.LG]