Title	**Learning Financial Time-Varying Networks**
Student name:	**Manuel Llamas**
Supervisor name:	**Mahesan Niranjan**

Aims/research question and Objectives

Background

Time-varying networks is recent field developed to explain and model the evolution of networks in different subjects aiming to understand how, when and why the networks evolve and even leading to predictive models of the phenomena tracked in the network. In network science distinct elements or actors are represented by nodes (or vertices) and the connections between the elements or actors as links (or edges) allowing us to study from how a single element can affect the whole network to even predict which topological changes indicate coming critical events or understand how a network can reshape to minimize contagion between its elements. All of this provides a powerful tool for studying financial networks, giving us a framework to track the evolution of the stock market and understand its behavior against critical events.

Hence the aim of this project is to study financial systems as dynamic networks focusing in the evaluation of the performance of some of the state-of-the-art existing methods in the financial field to find the network model that best learns and recovers the stock market behaviour, trying to build a more complete method by considering multi-node and multi-level networks.

Research objectives

- To find the most reliable model/s that can recover the financial market’s structure in different situations.
- Apply those models to recover and study the evolution of the stock market´s structure when facing different critical events.
- Study the relationship and effect of influential variables to the stock market like oil price, exchange rates or companies’ internal developments when recovering its structure facing those events.
- Upgrade the chosen algorithm to take into account multi-node and multi-level situations.

Summary of proposed research and analysis methodology

Basically we are going to focus on one method, TESLA (from the acronym TESLLOR, which stands for temporally smoothed *l*1-regularized logistic regression), however when reviewing the related literature they were found some other interesting approaches that could be useful to compare in terms of accuracy and efficiency.

TESLA

This algorithm is based on lasso-style sparse structure recovery technique and considers that temporally adjacent networks are likely to be topologically similar, therefore more likely to share common edges than temporally distant networks [1].
TESLA estimates the correlation (or dependency strength) matrix between the nodes using a time-series of the observed stated of the nodes $\mathbf{x}^{\mathbf{t}}$.

$$\hat{\theta}_i^1, \ldots, \hat{\theta}_i^T = \arg\min_{\hat{\theta}_i^1,\ldots,\hat{\theta}_i^T} \sum_{t=1}^{T} \ell(\boldsymbol{x}^t; \boldsymbol{\theta}_i^t) + \lambda_1 \sum_{t=1}^{T} \|\boldsymbol{\theta}_i^t\|_1 + \lambda_2 \sum_{t=2}^{T} \|\boldsymbol{\theta}_i^t - \boldsymbol{\theta}_i^{t-1}\|_1, \quad \forall i$$

The first term is the correlation estimator while the last two terms are the penalties ($L^{shrink} + L^{smooth}$) introduced to enforce sparsity and smoothness.
This algorithm had proved to recover dynamic networks in different fields as biology, citation and political networks, that makes it a great candidate to learn networks in further fields, there relies our interest on it.

TESLA updated

This update of the model uses, instead of the previous smoothing penalty (that considers small jumps between time steps) introduces a weighting function (ω_t^τ) that continuously smooth the changes in parameters [2].

$$\widehat{\boldsymbol{\theta}}_i^\tau = \min_{\boldsymbol{\theta}\in\mathbb{R}^{p-1}} \{\ell(\boldsymbol{\theta}_i; \mathcal{D}_n) + \lambda_1 \|\boldsymbol{\theta}_i^t\|_1\}$$

$$\ell(\boldsymbol{\theta}_i;\, \mathcal{D}_n) = -\sum_{t\in\tau_n} \omega_t^\tau\, \gamma(\boldsymbol{\theta}_i; \boldsymbol{x}^t)$$

The interest on studying this is to compare its efficiency with TESLA in the case that the time windows selected are very small (minutes), in such cases the change of the links in the network are expected not to change “abruptly”, as the original algorithm suggest.

Bayesian non-parametric model

This model was precisely used in financial studies [3] with feasible results, what makes it a great reference algorithm to compare the performance of the rest of them in the basic case already studied: the 2008 financial crisis.
However it assumes time-constant smoothness, while in finance we expect it to vary over time. This could be an interesting additional discussion when comparing its performance with TESLA, we expect the last one to recover networks in a more reliable way.

[1] Ahmed A and Xing E. (2009) Recovering time-varying networks of dependencies in social and biological studies. *PNAS* 106 (29) 11878-11883
[2] Kolar M, Song L, Ahmed A and Xing, Eric P. (2010) Estimating time-varying networks. *Annals Applied Statistics Vol. 4.* No. 1, 94-123
[3] Durante D. and Dunson D. (2014) Bayesian dynamic financial networks with time-varying predictors. *Statistics & Probability Letters*, 93, 19-26

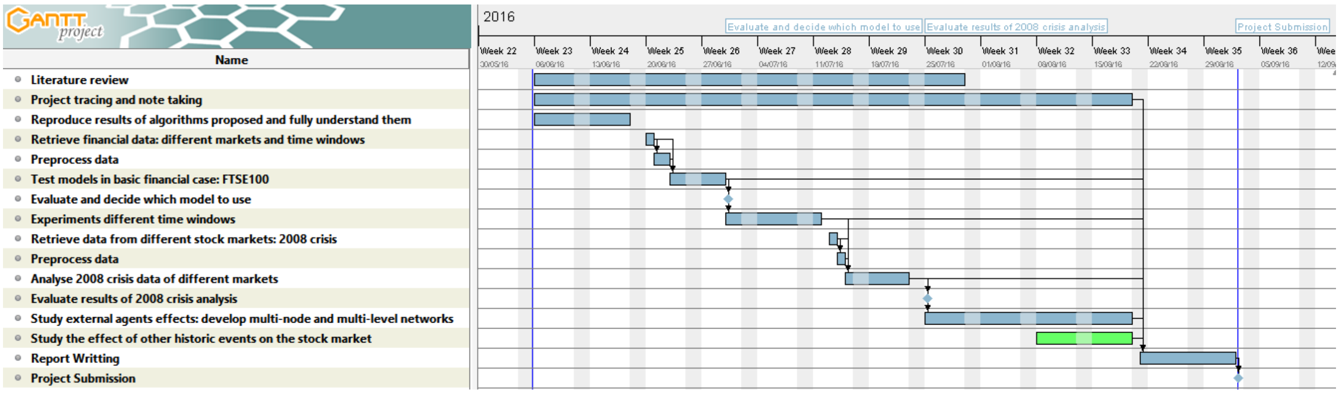
Research plan – Gantt chart or Pert chart

Knowing that the problem to treat is to test and extend existing Machine Learning algorithms the first steps comprise reproducing the results shown by the articles on where they were published to check they perform as the authors claim and understand the challenges they may have to face when recovering the studied networks.

Apply roughly the models choosing a standard time window and evaluate their performance comparing with previous work done in this field to evaluate how reliable are the different methods when applied to the financial field in order to select the model that best explains it.

After selecting the method/s to use, experiment different approaches to upgrade the algorithm/s: change time window (time steps of the dynamic network), analyse multi-node and multi-level networks and evaluation of the results.

Apply the resulting successful algorithm to study the effects of other critical events that strongly influenced the global stock market.



- Diamonds indicate milestones.
- Green color indicates optional task to develop if time allows it.

Ethical statement

In this proposed project all the data to analyse is public and don't involve personal information or privileged information about companies.

I, Manuel Llamas, student of University of Southampton declare that:

- I will attempt to achieve the best quality, effectiveness and dignity in the process of my work either the results obtained are correct or not.
- I will be honest regarding all my results. I will provide full information from all data, process and results obtained.
- I will avoid any harm that could make negative results that can affect the loss of information, property or social impact. I will take the responsibility if any harm is done.
- I will honour property rights.
- I will maintain confidentiality in case information come from a private organization or human being information unless I am explicitly allowed to do so.
- I will give credit and acknowledge every source of data and intellectual property to authors of works used in my research.

Legal and commercial aspects

The context in which this project is developed does not have any direct commercial propose. It is only focused on understanding how precisely financial networks can be recovered. The resulting product is not aimed to be used as a profitable tool, however if this work provides outstanding and reliable results could be further developed to predict market behaviours.
Summarizing the approach is closer to a research about the specific algorithms and their applicability in the financial area (apart from the already applied fields) rather than building a tool to get privileged information of future distribution of the market, therefore I don't consider possible commercial aspects at this early stage of the project.

Regarding legal aspects there is no reason to worry about since all data to use in this project is free to use.
In addition the project goal or techniques are not subject to especial regulations and therefore, don't require legal considerations.