Review of: "Multivariate Time-Series Data Generation in Generative Adversarial Networks"

Alessio Staffini

Potential competing interests: No potential competing interests to declare.

Dear Authors,
I am glad to have had the opportunity to read your work. The study might be interesting, but I believe many aspects should be dealt with before considering the manuscript to be ready for publication. Please see such points below, divided in major and minor comments.

Major comments
---

1. English writing can be highly improved. As presented now, many parts of the paper are very hard to follow, and the meaning of some sentences is often unclear. I strongly suggest to send the manuscript to a proofreading company.
2. Please also define the various acronyms appropriately when they are first mentioned, reorganize the sections to have a more coherent flow, and correct typos in the text (for example, in Background Knowledge you mention that 0 represents fake data with probability 0, but you probably wanted to say that 0 represents fake data with probability 1). At the moment, there is a lot of "confusion" (on the goal of the work, on how the empirical analysis was conducted, and on the obtained results).
3. It is unclear to me what is the main contribution of the manuscript. You mention in the abstract that you generate realistic time series data using only a vector of noise, but this is the standard GAN approach for generative tasks (while I agree that for time series forecasting what is usually provided is not noise but a matrix of predictors, forecasting does not seem to be the focus of this paper). For example, reference [2] in your manuscript (Takahashi S., Chen Y., and Tanaka-Ishii K., "Modeling financial time-series with generative adversarial networks") already does what you describe as novel for financial time series data.
4. In Results and Discussion, if you argue that data augmentation using your GAN improves the performance of classification tasks, you should show the performance of the models also before adding the generated data. More details on the experiment (ex. what is the dimensionality after PCA?) are required.

Minor comments
---

5. GANs are known to be unsupervised learning models, but the discriminator is trained in a supervised fashion. In this work you still require real samples for training, so sentences like "we adopt a novel approach that is completely unsupervised" are unclear to me.
6. The literature review could be expanded by including other related works. At present, there are only 13 references in

total.

7. The manuscript would benefit from the addition of figures that show how generated data compare to the real samples.