Geoscientific Model Development Discussions

Geosci. Model Dev. Discuss., community comment CC1
https://doi.org/10.5194/gmd-2021-402-CC1, 2022

Model comparison and naming conventions

Oscar Jacquot

Community comment on "Conservation laws in a neural network architecture: enforcing the atom balance of a Julia-based photochemical model (v0.2.0)" by Patrick Obin Sturm and Anthony S. Wexler, Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2021-402-CC1, 2022

Dear authors,

I would like to make a few comments comments about your paper :

1) This comment concerns the choice of your baseline model and subsequently the way some results regarding your approch are assessed.

In section 2 you introduce two networks. One is a simple two layers dense neural network which you refer to as naïve. The other network, which is the focus of your work, features two architecture modifications with regards to the naïve network. One affects the input with the addition of bilinear terms correponding to the two bodies reactions in your chemical system. Another modification affects the output, resulting in your network effectivelly predicting the rate of each reaction in the second to last layer and subsequently assesing the change in concentration of each species in the last layer, according to stochiometric constraints.

In section 3.4, you compare carbon and nitrogen conservation between the two networks, effectively showing that your stochiometric constraint is indeed verified while the naïve implementation fails to do so. This is indeed a predictable property from your network, as this constraint is in some sense hard coded and is independant from the ability of your network to predict reaction rates well.

Yet, one could reasonnably assume that providing the naïve network with relevant bilinear terms as an input without including hard coded stochiometric information would on its own improve forecast performance. Therefore, I believe that your paper currently has a blind spot : what are the relative contributions of the the input modification and the stochiometric constraint to the improvement of forecasting ability with regards to the naïve network ?

I think that your paper would benefit from displaying the results obtained with a third intermediate network, which would include the input modification but no stochiometric constraints on the output. The comparison of correlation and element conservation for all 3 networks would be more insightful, but should not require much work as the same learning procedure can be applied for the intermediate network.

2) This comment concerns naming conventions.

You chose to refer to your network as a physics constrained neural network. This name sounds similar to the one chosen by Raissi et al. (who you refer to), regarding physics informed neural networks (PINNs) which are now a widely shared name and concept in the community.

In your introduction, you clearly state that PINNs are neural networks which satisfy a given partial differential equation (PDE). Specifically, the physical information is carried via the PDE and enforced in the cost function, and the only inputs to a PINN are the variables that you differenciate with regards to in the PDE. Clearly, your network does not satisfy such a definition and you do not claim so.

Yet there are several occurences where your wording does not carry that distinction so clearly, or could make the reader confused. More preciselly, there are several occurences of the phrases "physics-informed input" or "physically informed input", such as in lines 116, 126, 130, 138 and 491. Refering to the addition of relevant bilinear terms as an input to your network in such a way is extremely confusing with regards to the widely shared definition of what a PINN represents.

Even though calling your network a phyiscs constrained neural network seems justified and different enough to me, I believe that you should rather refer to your input modification as a "reactivity-informed" input, or any other better suited name of your choice. Currently, I believe that your naming convention could generate a lot of confusion between widely different implementations.

I also believe that the reference to PINNs regarding the addition of the stochiometric matrix A in line 102 is unsupported.

3) On line 149, you reference a Zenodo archive which is supposed to redirect towards a Julia source code used to generate your data. However, I find that this archive actually hosts a publication by Gabay, A., Duhamel, P., & Rioul, O. related to signal analysis. Could you correct this link, as there seems to be a mistake ?

Best regards,
Oscar Jacquot