Comments to the Author   
Building on previous works that point at the influence of iron in the ideality factor of silicon solar cells, the authors present a deep neural network that can predict the iron content, and explore its range of validity with simulations.   
  
The approach is interesting, and opens a path for further improvement and exploration of machine learning techniques to extract physic parameters from electrical measurements. Its steps are well justified, but in my opinion the discussion of results needs further clarification.   
  
For example:   
-What does it mean “Which unused during network learning (B-varied dataset, Fig. 4(c))”? The DNN is not trained with 9 NB values from 1e15 to 1e17?   
-I see a lack of rigor in the sentence “But the expected improvement of the nFe−FeB–nFe DNN must be also caused by the removal of a certain degeneration of correlation between an ideality factor value and an iron concentration (a kind of splitting).” A “certain degeneration”? A kind of splitting?   
-“At the same time the increase in number of samples leds to essential raise trainability of network, which uses two ideality factor value.” I am sorry I do not understand what is meant here.   
  
Another point that needs further elaboration is the fact that the model of the I-V curve in which the DNN is based does not take into account resistive effects, which tend to be present in real solar cells. I am aware that incorporating this issue is beyond current work and a matter of future research, but I would like a comment in the discussion on how would the presence of the resistance effects influence the DNN results.   
  
A couple of suggestions about the figures:   
  
-Is Figure 2 really informative? I am not so sure, it is very colourful but also so tupid that it is difficult to visually understand what is going on…   
-I would also recommend to better distinguish the nFe−FeB DNN and nFe−FeB–nFe DNN cases in Figure 4. Separating it in two figures is an option, or labelling the figures and not only indicating it in the figure caption.   
  
Finally, I have to say that the paper needs a thorough revision of English, preferably by a native or bilingual speaker. English is not my mother tongue, but I think that there are many expressions that are not correct, and make the reading difficult. From the title itself “An evaluation for?”, and much more. Some examples: “a not little collection (…) an (…), an (…), a…”, “which determined for experimental data”, “The work milestones are following”, “to following silicon properties”, “The pair’s fraction does not constant in SC regions“, “n can takes”, “In practice terms, the obtaining of…”, “which used for…”, “expected to enhances”, “The enjoyable (from a practical point of view) attribute is an ability to predict iron concentration value, which not used under learning.”, “does not require a much experimental time”.

The manuscript "An Evaluation for Iron Contamination in Silicon Solar Cell   
Using Ideality Factor and Machine Learning" by Oleg Olikh, et al, reports prediction of iron concentration using deep neural network. However, the utility of models for predictions has to be built upon calibrations of the model to existing experimental data and getting estimates of uncertainties. If not and unfortunately it is the case for the manuscript, the model should be considered far from being validated.     
I suggest the authors considering analyzing their own fabricated solar cells and/or existing reported solar cells in the literature to prove the validation of the DNN model to improve the quality of manuscript. In addition, there is plenty of room for improvement in English expressions.

The work seems to be reasonably carefully done, but is not of especially high impact. In addition, the English is poor, in places poor enough to compromise clarity (especially in the end of the discussion section on page 4). I believe it would be more appropriate as a Brief than as a full article, as there isn’t sufficient technical analysis/depth. To shorten the published length sufficiently, I would suggest removing Figures 1 and 2 and/or moving them to SI, as their content can be conveyed through the text.   
  
More specific technical critiques:   
  
1. The explanation of the different datasets (last paragraph of Section IIC) is unclear, and Figure 2 is very difficult to interpret. I don’t understand what the open vs. closed symbols are (perhaps train vs. test?), nor why some of the points in the grid are not evenly spaced. This data would likely be better presented as a table with lists of grid coordinates associated with each subset.   
  
2.The neglect of series and shunt resistance in Equation 2 is not justified.   
  
3. There should be some comparison of the precision obtainable from the model presented to that achievable from typical experimental methods.   
  
4. There should also be some discussion of why the errors are so much larger on the B-varied test set. I think this may be what is addressed by the paragraph beginning at line 22 on page 4, but I can’t really interpret what that paragraph is saying.   
  
5. If the authors really wish to claim that this approach could be useful on a manufacturing line, it would also be good see a comparison of this approach with a conventional interpolation+regression approach for the same features: e.g. use the same simulation data and create hypersurfaces (analogous to those in Figure 3), then predict N\_Fe by simply interpolating on those surfaces, which could be more robust to extrapolation and should be better able to detect outliers.