Dear editor

Thank you for the referee report. The authors have studied the reports of the two reviewers and would like to ask you to reconsider your decision.

Both reviewers wrote very positive reports of the paper praising both the impact and the quality of the work.

The criticism of Reviewer 1 due to a previous purely theoretical study about multi-task-learning neural network architectures is understandable but not founded. The earlier work was an analysis of neural network architectures (feed-forward, multi-tasking, and so on) to identify with generated data which network architecture is best suited for this type of problem (multi-parameter regression). The new paper submitted to Optica uses the architecture, which was previously found to work best. However, the work which we would like to publish is not that. First of all, the paper which we submitted to Optica is on a real physical optical sensor that was built, characterized, and used for acquisition of data for the training and the test of the novel approach. Secondly, the paper submitted here for publication does much more. It defines how the data acquisition should be carried out, which is essential to a sensor based on machine learning.

Furthermore, it introduces a new metric, the Error Limited Accuracy, which bridges the gap between two fields: physics and computer sciences. We strongly believe this metric is necessary for the characterization and, thus, the spreading of a new generation of sensors into applications. In conclusion, this work demonstrates that a new generation of sensor which are based on the proposed approach is not only possible but would work better than the conventional ones. This work thus represents a paradigm shift. This is the reason why we strongly think the work is of great relevance for the public of Optica and for the optical community in general and should be published on Optica.

We are willing to improve the text to address the issue pointed out from the first reviewer and the minor points of the second reviewer.

Since both authors recommend the paper for publication, we strongly ask you to reconsider your decision.

Best regards

----------------------------

Reviewer 1:

The authors presented a multi-task learning neural network model to extract oxygen concentration and temperature by evaluating the dynamical luminescence quenching by oxygen molecules, or as the authors named, “a learning sensor with parallel inference”. This manuscript is well presented, the model could work effectively and correctly as to my knowledge. However, formerly on Applied Science, the authors already has a publication (Ref. 26), which described how they obtained the neural network parameter of this model. This largely undermined the contributions of this manuscript. Could the authors state the novelty of this manuscript from their former publication? If not, I would recommend this manuscript not to be published this time.

Reviewer 2:

The authors did a great job in writing the paper. The paper is well written, and the method described in the article is interesting. The main contribution of the paper is in the use of a neural network model (NNM), namely Multi-Task Neural Network, to model the measurements of oxygen and temperature based on luminescence quenching. The model predicts these two measurements based on one set of optical measurements. Instead of synthetic data as in authors previous work, the authors trained the NNM based on real experiments data. An automatic data collection method was used to collect 189,000 training data.

However, several minor points can be improved, as follows:

1. In section C.1, the authors described two types of input that were used to train the network. Please explain why the phase shifts of the first input (theta\_s) are divided by 90.

2. Is there any particular reason why the mini-batch of 32 samples was used?

3. Are the results in the Sensor Performance Section from the training data set or the testing data set (new data that never been seen by the model)? If it is from the testing data set, could the authors describe its data collection?

4. To generate the model, it requires 65 hours of training data collection and an additional 5 hours for training the model. Could the authors discuss whether this process should be repeated every time a new sensor is manufactured? Or is there any way to translate the generated model into a new sensor?