The paper entitled: ***SIMPOR: Synthetic Information towards Maximum Posterior Ratio for deep learning on Imbalanced Data****.*

By: Hung Nguyen and J. Morris Chang

Dear Editor,

We hereby submit our manuscript entitled ‘***SIMPOR: Synthetic Information towards Maximum Posterior Ratio for deep learning on Imbalanced Data****’* to the IEEE Transactions on Artificial Intelligence.

This work studies the effects of class imbalance mitigations on deep learning models by providing several experimental results and visualizations. While most previous studies focus on deep learning with the image data type, we aim to address the issue of tabular data as it is crucial but missing in the field. We also propose a data balancing algorithm to mitigate the negative impact of class imbalance by generating synthetic data for the minority class. Our proposed algorithm leverages a combination of a heuristic method to find where should focus on generating synthetic data (high entropy region) and a conventional statistical theorem to determine the direction to spawn synthetic data. Since our technique is directedly designed for deep learning, it is shown in the experiments that it outperforms other state-of-the-art techniques in terms of deep learning performance (F1-score and AUC). Our experiments with 41 real datasets show that the proposed technique achieved the highest winning times compared to state-of-the-art techniques. We also provide several data visualizations of each technique to understand how techniques perform. The visualization shows that other techniques can accidentally create many noisy samples because of poor-placed synthetic data, while this number of the proposed technique is very small.

With the intuitive and promising results on 41 real datasets shown in this manuscript, we hope that the community will be beneficial from our study. In the end, we should say, the manuscript has not been published or copyrighted in any formal or informal journal.

Sincerely,  
Hung Nguyen

Ph.D. Student

#EE, University of South Florida,  
Tampa, Florida, 33617 USA.  
Phone: +1-8137667064

Email:  nsh@usf.edu