Referee Report on the paper entitled "Some Universal Insights on Divergences for Statistics, Machine Learning and Artificial Intelligence"

The paper/chapter under review presents a universal framework for divergence measures motivated by their enormous applications in statistics, machine learning and artificial intelligence. The authors have also presented the general formulation of divergence between two different types of functions, which appears to be a very useful tool. It is a very interesting and useful piece of work along the lines of the current state-of-the-art of the field. Here are some minor suggestions to improve the presentation of the paper for more general audiences which the authors may consider if they feel appropriate.

1. Section 2 contains several ideas and concepts from different fields where the divergence measures are applied. The literature survey of this section is extremely comprehensive. However, the presentation assumes that the reader is reasonably familiar with the area. To reach the readers who are not familiar with some of these ideas, a presentation in a somewhat simpler language could have been more useful. Of course all depends on the authors and the editor, and if the current set up is what the authors intended then that is fine. After all, it is not as if the particular segment is deficient in any way.

2. Section 3 onwards, the paper is extensively rich in theory and hence there are a lot of notations. Frankly, I found myself lost among the notations after a few pages. I think, it would be much easier for a reader to follow the main ideas and results if they can be separated out from the notational worry. The authors could add a subsection at the beginning of Section 3 and list there all the notations to be used throughout the rest of the paper. Then, they need not to define these notations in-between text and the presentation will be much clearer. Also, it will help readers to consult this list of notation (at one place) whenever needed.

3. For the same reason as above, I suggest that the authors move all the proofs to a separate section at the last or to an appendix.

4. It will better to write "σ-field", "x-th", "λ-a.a.", etc. in place of "$\sigma-$field", "$x-$th", "$\lambda-$a.a.", etc., respectively.

5. Section 4 contains different ways of solving minimum divergence problems between two different types of functions. If the authors provided some numerical comparisons in this section

that might have been helpful. This actually is a relevant comment for other sections of the article also. The article is entirely theoretical and contains no data.

6. In some cases, existing words in English may have been extended to form new words, but I am not sure if they are legitimate English Words. E.g., “flexibilizing”, or “abbreviatingly;.