Machine Learning Requirements and Guidelines for the Journal Additive Manufacturing

To assist both authors and reviewers, and to ensure a consistent evaluation of article expectations with regard to machine learning (ML) submissions to Additive Manufacturing, we now ask that all submissions using ML consider the following rubric within the context of the specific problem/algorithm(s) at hand. Authors may use their own discretion as to whether the description/discussion of each point of the rubric belongs in the main text, or if some points are better suited for an appendix or online supplement.

Elements of a Machine Learning Research Article in Additive Manufacturing:

1. Is a clear description of the dataset provided? What are the inputs? what are the outputs? how was the data generated and/or aggregated? Can the scientific community access the dataset for future work and/or to verify the results (not a requirement, but should be addressed)? If so, how? Is the dataset sparse or dense? If relevant, what techniques were used to handle missing values? Where any feature engineering approaches used to manipulate the raw data sources? If so, are they adequately described and documented? Were new feature engineering approaches developed? If so, are they clearly explained and is their impact to the scientific community put into proper context (are they bespoke to this work or could they be adopted in solving other problems)?

2. Is the underlying AM problem well-defined for machine learning? Machine learning, like finite element simulations or electron microscopy, is a tool available to help solve or understand a problem, but not an applicable tool for every problem. The details of how ML is brought to bear to solve a problem can be instructive in guiding how other problems are approached, but the context and scope of the problem will impact the choice of choosing to use ML at all, and then again in choosing the best ML algorithm. Have the authors adequately explained why and demonstrated that ML is an appropriate tool for the problem at hand? Does a basic statistical assessment (visualization, correlation analysis, etc.) of the data that is to be modeled using machine learning provide answers to the following questions: how is the data statistically distributed? How strong are the statistical correlations between the inputs and outputs that are going to be modeled and/or do correlations exist? Is ML feasible for solving this problem, or are the data points of interest statistical outliers relative to the best statistical descriptions/models of the dataset?

3. Is ML necessary? Once it is determined that the data of interest exhibit statistics that warrant statistical modeling, the next questions that should be answered is: Is ML needed, or do basic statistical tests elucidate the statistical correlations statistically in and of themselves? i.e., are the some of the correlations required to solve the problem/answer the question of higher dimension than we can determine without the use of ML algorithms?

4. How do the basic statistical properties of the dataset motivate the choice and selection of machine learning approaches/algorithms? Is there adequate reasoning/rigor provided in stating why the
algorithms that are being considered are the best choices to evaluate for this specific dataset or problem? If there is not an obvious scientific/mathematical reason to choose one vs. another, have the authors evaluated all feasible algorithms for their problem and used appropriate methods and metrics to select the best performer(s)?

5. Is the tuning of the ML algorithm/model(s) well-reasoned and documented? Do the authors address the implicit limitations of the models, such as the independence condition in Naïve Bayes or gradient collapse in convolutional neural networks, and the impact these limitations have on the applicability of a particular model? Were proper model optimization steps taken in tuning all hyperparameters? Is the bias-variance tradeoff discussed? Are the full range of hyperparameters reported? Does the article provide sufficient instruction for others to identify the decision points and considerations necessary to construct a similar class of model on an unrelated data set? Can those models be repeated by others? Has the code been packaged properly to ensure compatible versions of the necessary packages, e.g. tensor flow, keras, scikit-learn, numpy, scipy, Matlab, etc.?

6. Are the ML model performances properly assessed? Are bias, precision, and variance considered in the uncertainty quantification (i.e., are more than just mean value error assessment(s) provided)? Have the models demonstrated robustness to both extrinsic vs. intrinsic error (i.e., are they robust against both overfitting and underfitting)? Was the dataset properly divided into training vs. test data, and did the authors properly test their models?

7. Is scientific and/or engineering impact clearly derived from the use of ML? Did the models solve the problem at hand? Did the models perform better than random guessing and/or the previous state of the art without ML? Did the authors properly quantify how well the ML algorithms work for their problem? Does the performance of the ML models provide evident scientific or engineering impact? How was this impact quantified?