Dear Dr Wang,

We have received the decision on your paper entitled "A Novel Robust Data Synthesis Method based on Feature Subspace Interpolation to Optimize Samples with Unknown Noise".

The Editor's and the reviewers' comments are as follows:
****

EiC
A very thorough editing is required; consult a professional and show all changes made to the revised manuscript.
References: max 50 entries. Make sure the references are recent and highly pertinent to the contents of the manuscript.

Reviewer #2: 1.The authors examined data synthesis method based on feature subspace interpolation to optimize samples with unknown noise.
A novel data synthesis approach (the Robust Subspace Interpolation Synthesis - RSIS) is proposed.

2.The suggested approach consists of four stages (four-stage composite solving framework):
Stage 1. The unsupervised clustering algorithm (the original dataset is divided into multiple subsets with approximately equal sample sizes, each corresponding to a feature subspace).
Stage 2. The concept of the Traveling Salesman Problem (TSP) to order the feature subspaces.
Stage 3. Integration of the mechanism of soft parameter sharing, conducting a linear fitting between adjacent subspaces considering global information based on the ordering results.
Stage 4. An innovative multi-stage minimum weight matching method, through which an suboptimal interpolation matching strategy is obtained.

3.The article material on experiments are presented at a very good level.

4.In general the article can be accepted:
(i) the topic of the article is very important;
(ii) the suggested method (RSIS) is very interesting and prospective;
(iii) future research described;
(iv) all parts of the article materials are presented at a good level.
The article material will be useful for many readers.

Reviewer #3: All comments are well addresses.

Reviewer #4: Summary:
The authors addressed most of my raised concerns. Some of them were replied to me in the Revision Note, but they did not add them to the paper. Therefore, my minor suggestions, as described below, are that some of these replies could be useful in the paper either to increase reproducibility or the reader's understanding.

Minor Suggestions
1. For readers trying to reproduce the proposed method, the details of the LOF algorithm are important. As explained by the authors in the review reply, the number of neighbors parameter for LOF is not critical. Perhaps a sentence would be sufficient to increase the reproducibility of the proposed method. Something as "The LOF method primarily involves the hyperparameter k (number of neighbors), which represents the number of neighbors used for calculating the local density of each data point, i.e., the number of points in the k-distance neighborhood. In practical applications, the setting of this hyperparameter is not particularly critical and does not significantly impact the effectiveness of the RSIS method.". In this way, readers who attempt to reproduce the method would understand not to focus on the LOF hyperparameters.

2. I appreciate the explanation of the kNN optimization given in the Revision Note. If I were a reader reading this paper for the first time, I would still be intrigued as to why the kNN optimization works. I know the paper is already dense and this part is not the focus, but maybe part of the explanation given to me could be included in the paper to satisfy the reader's curiosity in this matter. The excerpt concerning the goal of the clustering step ("Our purpose in clustering is to meet the assumptions of the RSIS method as much as possible and to ensure a good connection with other algorithms of RSIS (such as multi-stage minimum weight matching)") is especially enlightening, as it also presents that the clustering technique does not need to be optimal; it only needs to be good enough to help the next step.

3. I understood the point about Fig. 9 starting point. I still find it very confusing because the metrics are not comparable; thus, there is no readability gain to keep them in the same chart. The paper is comparing the three configurations using p1 and p2. I suggest separating the metrics into two charts or explaining the 0.08 starting point in the text as explained in the Revision Note.

4. Because it is a new complex method, comparing it to other methods is not trivial or direct. Perhaps in the Conclusion section, the authors could point to a future direction where the advantages of using this method could be compared to other methods.