

Dear Editor and Reviewers,

According to your suggestions, we have prepared a new version of our paper, introducing the changes required to incorporate all the reviewers' comments. The following is our explanation of the changes made, as well as how they address the comments raised by the reviewers. Text in bold corresponds to the original comments. We hope to satisfy the reviewers' expectations. Finally, we are really grateful to the reviewers for their insightful and constructive comments and suggestions in their revisions, which have helped us to improve the quality of the article.

We hope that the reviewers find the proposed thorough revision of our manuscript satisfactory. We look forward to hearing from you soon.

1 Reviewer A

Although this paper presents three multi-target support vector regression models, I feel that the paper has not sufficient contribution to publish in INS, based on the following concerns.

We are grateful to the reviewer for their insightful and helpful comments. They have allowed us to improve the paper and enhance the scientific merit of the work.

We have conducted a thorough revision of the paper to highlight the scientific contribution, novelty, and performance of the methods proposed in this paper. Specifically, we have stressed the main contributions by editing and including the following bullets in the introduction of the new version of the paper:

- *Evaluating the performance of a Support Vector Regressor (SVR) as a multi-target to single-target problem transformation method to determine whether it outperforms current state-of-the-art Single-Target (ST) algorithms. We analyze its performance as a base-line model for MT chaining methods due to the fact that ST methods do not account for any correlation among the target variables.*
- *Building a Multi-Target (MT) ensemble of randomly chained SVR models (SVRRC), an algorithm adaptation approach, inspired by the state-of-the-art chaining classification method, Ensemble of Random Chains Corrected (ERCC), to investigate the effects and advantages of exploiting correlations among target variables during model training, in the context of regression problems. The main issues to be investigated with this approach are the randomness of the created chains because they might not capture of correlations between the targets, as well as the time taken to build all the regressors in the ensemble.*
- *Proposing an MT algorithm adaptation model of SVRs that builds a unique chain, capturing the maximum correlation among target outputs, named SVR Correlation Chains (SVRCC). The advantages of using this maximum correlation chain approach include exploiting the correlations among the targets which leads to an improvement in model prediction performance, and a reduction in computational complexity because a single SVR-chain model is trained, rather than building an ensemble of 10 base regressors.*

These contributions advances the field of multi-target regression by offering more accurate and less computationally demanding models.

1. It seems a simple collection of simple methods. What is the respective advantage for each model?

The answer to the previous comment addressed the main contributions of the three models. Specific similarities and differences among the methods are detailed in Section 3, which we have rewritten to clarify and stress them. Specifically, we included three new diagrams to help to visualize the differences and improved their description in the text. In short, the differences, advantages, and novelty of the models are the following. The first model evaluates the performance of SVRs using a MT to ST problem transformation. This is useful to

analyze how good are SVRs in contrast to other base regression methods, but does not take into account at all any correlation among targets. The second model evaluates the performance of an ensemble of random chains using SVRs, inspired in the best performing algorithm in the state-of-the-art (ERCC). The main problem of this approach is the *randomness* of the chains employed and the significantly larger amount of time required to build all the base regressors in the ensemble. The random chains approach does not guarantee to exploit correlations among targets in the chain. Finally, the third model builds the unique chain with maximum correlation among targets and then learns a single SVR-chain model. The advantages of this approach are: 1) exploiting the correlations among targets lead to better results according to the experimental study, and 2) only requires to learn a single SVR-chain model, in contrast to the ensemble approach that builds 10 base regressors, thus it's significantly faster, as also shown in the experimental study. The use of maximum correlation chains has not been addressed before in the context of multi-target regression and the experimental study supports its better performance. Therefore, we're proposing a novel method which is both faster and obtains better results. Interestingly, the state-of-the-art method ERCC proves to perform as the slowest method.

2. The difference and relationship among the three models have not been addressed well.

As indicated in the answer to the previous comment we have rewritten Section 3 to improve the description of each of the models and their differences. Moreover, we included three diagrams to show the flowchart of the output predictions, we edited the algorithms' pseudo-codes, and remarked their similarities and differences in the text description.

3. Analysis on the algorithms is obviously insufficient.

We improved the description of the algorithms in Section 3, as indicated in the previous comments, to enhance and strengthen the theoretical foundations of SVR learners. We added a new section to the paper in order to analyze the strengths and weaknesses of each model, identify their areas of applicability, and formulated suggestions for the readers on when and which model to use. Moreover, we have significantly extended the experimental analysis as indicated in comment #5.

4. I cannot see the key advantages of the three models and their performance.

Advantages of the models, addressed in previous comments, specifically in the bullet points in comment #1, have been clarified both in the introduction, description, and conclusions of the new version of the paper, as well as adding a new section to discuss the results and highlight the models' advantages. Their performance is analyzed throughout the extensive experimental study, achieving better and faster results (considering 4 performance metrics and runtime) than the other algorithms compared, especially when compared with ERCC, the state-of-the-art method which is shown to run the slowest.

5. Simulations and validations are not enough.

We have significantly extended the experimental study in many ways:

- Included three new performance metrics (average correlation coefficient, mean squared error, average root mean squared error), together with the previously existing average relative root mean squared error. By using many metrics, we capture different and complementary behaviors of the methods on the datasets, offering a diverse point of view on their performance.
- Included 6 new datasets making a total of 24 datasets with a varied number of instances, attributes, and output targets.
- Included a new algorithm (Regression Chain) making a total of 10 algorithms compared.
- Included a runtime comparison of all the algorithms to analyze their scalability as for the number of instances, attributes, and output targets.
- Included a new statistical analysis considering the new performance metrics and runtime comparison on the Friedman, Bonferroni-Dunn, Nemenyi, Holm, and Wilcoxon tests.

We thank the reviewer for his comment on extending the simulations and validations study. We believe we now have a thorough and comprehensive experimental study that emphasizes the better performance of our contributions when compared to the state-of-the-art over many performance metrics along with an extended statistical analysis.

6. Conclusions are neither clear nor concise.

We have rewritten the conclusions to clarify and improve our contribution and insight of the results obtained, and concisely edited the concluding remarks. Thank you.

2 Reviewer #3

This paper proposed the Support Vector Regression via Correlation Regressor Chains. This study mainly process data with random chains and maximum correlation chain. Some sincerely comments as follows, for the author(s), may be useful to improve your work:

We do appreciate your sincere comments about the article and the scientific contribution. We believe the new version of the article has been significantly improved thanks to your insightful and comprehensive comments.

Comment 1. This study used the SVR based to multi-target prediction. In introduction section, the manuscript should investigate the SVR. Why does the research adopt the SVR? Furthermore, have the random chains and maximum correlation chain been examined in any literatures?

We have rewritten the introduction and background to include the SVRs research to highlight their importance, advantages, performance, and scalability, in order to justify our approach on why using SVRs for the multi-target regression.

There are many studies in the literature on the use of random chains in multi-target regression. Indeed, the best performing state-of-the-art method is Ensemble of Regressor Chains Corrected (ERCC), which builds an ensemble of up to 10 random chains, and it is included in the experimental comparison. However, there are not many studies on the exploit of correlations among output targets. Zhang et al. presented a multi-target SVR approach that takes into account the correlations between all the targets using the vector virtualization method. It extends the original feature space and expresses the multi-output problem as an equivalent single-output problem, so that it can then be solved using the single-output least squares support vector regression machines (LS-SVR) algorithm. Nevertheless, to the best of our knowledge, our work is the first to conduct an study on the use of the maximum correlation chain in multi-target regression problems. We have rewritten the background section to clarify these studies in the literature.

Comment 2. Reviewer suggests that the flowchart of model should provide in section 3.

We added three new flowcharts, one for each method, in Section 3 to clarify the description of the models and their differences.

Comment 3. The results should provide the parameters of SVR. Many literatures have proved that the parameters will influence the performance of SVR.

The reviewer is totally correct about the influence of parameters in SVRs/SVMs. Therefore, we clarified the parameter settings of the algorithms in Section 4.2 and how the best parameters are selected via cross-validation. Cross-validation conducts a sensitivity analysis on the combinations of hyper-parameters, ensuring the best parameters are selected for each dataset.

Comment 4. Reviewer confuses that some datasets have not be applied in regression field which are for cluster or classification. Author(s) also should investigate the best performance in current literatures.

The datasets were collected from the UCI Machine Learning Repository and the Mulan website for multi-label learning. It is true that many versions of them are available both for classification, regression, or clustering having the same dataset name, but their outputs are specific to the learning problem. The datasets we employ are available publicly online in these dataset repositories and are specific for multi-output regression. Their characteristics (especially the number of output numeric targets) are detailed in Table 3. The experimental study compares with 7 other state-of-the-art methods in multi-target regression literature.

Remarks: On page 5, “Multiple authors have proposed Multiple-Target Support Vector Regression methods [5, 43, 44]”. “Multiple authors” should be changed to “Many authors”.

Changed. Thank you.