October 5, 2018

To the Editor:

Please find enclosed the revised version of our jointly authored paper “Inducing Non-Orthogonal and Non-Linear Decision Boundaries in Decision Trees via Interactive Basis Functions” that you kindly invited us to revise and resubmit for publication consideration to Expert Systems with Applications.

First of all, we would like to thank two anonymous reviewers for their feedback on the first version of this paper. We believe that their encouragement to expand our experiments has led to a much stronger paper. More concretely, whereas the previous version of the paper was based on less than ten datasets, now the benchmarking experiments are conducted more than ninety publicly available datasets. In addition to addressing the issue of bias in data set selection for the experiments (a concern raised by Reviewer #2), this also provided a more voluminous set of results that could be systematically proved. In this way, we are now able to provide more concrete guidance with respect to the conditions under which Interactive Basis Functions tend to perform better than orthogonal partitions for a given algorithm.

Another change that is worthwhile mentioning is that given the size of the experiments in the revised paper, we dropped the time benchmarking part, as benchmarking execution time would have been impractical (in our estimation, given our resources, benchmarking @ 50 repetitions of each algorithm would have taken no less than two months of computer time).

In this letter we describe all the revisions made in response to the feedback received. We trust that you will find that our actions provide a meticulous and conscientious response to the recommendations from the reviewers.

We thank you again for the attention paid to this submission and look forward to hearing back from you in due course.

Sincerely,

Antonio Páez

(on behalf of the coauthors)

**Reviewer #1 (responses in blue)**

This paper presents a technique of inducing non-orthogonal and non-linear decision boundaries in decision trees. Although the core idea is not new, this paper intuitively employs Radial Basis Functions for this purpose.

Thank you for your comments and suggestions, which have helped to improve the quality of the paper.

This paper has the following major drawbacks.

The experiment section of this paper is inconclusive as rigorous evaluation on classification accuracy is missing. The authors should compare classification accuracy with other decision tree algorithms that are employing linear/orthogonal splits on a number of (30 or more) widely used, publicly available data sets preferably from the UCI Machine Learning Repository. As non-orthogonal and/or non-linear splits compromise knowledge interpretability and are more computationally intensive, the authors should concretely establish their superiority theoretically, experimentally and statistically (over iterative linear/orthogonal splits).

Many thanks for this suggestion (also see comments of Reviewer #2). We have availed ourselves of a subset of the collection of data sets provided by Fernandez-Delgado et al. (2014). Accordingly, whereas in the original paper we used only a small number of data sets for the experiments (<10), now the benchmarking experiments are based on 93 data sets. Our results are in line with the results of Fernandez-Delgado et al. (who find that random forest is generally the best classifier), add one classifier that Fernandez-Delgado et al. did not consider (evolutionary tree), and indicate in general the conditions under which IBFs tend to provide better performance. Furthermore, we introduce a device called Decision Charts that helps to improve the interpretability of DTs with non-orthogonal/non-linear partitions.

Also, the authors should discuss more on SVM and put SVM in comparison spectrum.

We introduced additional references. However, we decided against putting SVM in the comparison spectrum. Fernandez-Delgado et al. already show that SVM performs better than some implementations of DTs, but not random forests. Rather, our objective here is to compare IBFs within DT algorithms. We hope that you will agree this is a reasonable way to proceed with the experiments.

The authors should computational information in the experiment section.

Done.

Figures should be positioned in the text.

Done.

**Reviewer #2**

The paper augments random forest with IBF. The paper is well written and easy to follow.

Many thanks for your constructive comments.

However, this reviewer has the main concerns regarding the experiments parts. The authors only conduct experiments on very limited number of dataset and compared with vey limited number of baselines. Recently the researchers are using the following JMLR benchmarking to conduct large-scale comparison to remove the bias in dataset selection

a) "Do we need hundreds of classifiers to solve real world classification problems?."

Many thanks for this suggestion (also see comments of Reviewer #1). We are particularly grateful to you for bringing the paper of Fernandez-Delgado (2014) to our attention. Thanks to it we were able to substantially improve the quality of our experiments. More concretely, we have availed ourselves of a subset of the collection of data sets provided by Fernandez-Delgado et al. to increase the number of data sets for the experiments from less than 10 in the original paper, to more than 90 in this revision. In general terms, our results are in line with those of Fernandez-Delgado et al. (who find that random forest is generally the best classifier). We also incorporated one classifier that Fernandez-Delgado et al. did not consider (evolutionary tree). And a meta-analysis of the benchmarking results provided very informative results regarding the conditions under which IBFs tend to provide better performance for a given algorithm, relative to orthogonal partitions.

Several important reference on oblique random forest are missing:

b) On oblique Random Forests

b) Random Forest with ensemble of feature spaces

c) Oblique Decision Tree Ensemble via Multisurface Proximal Support Vector Machine

d) Robust visual tracking using oblique random forests

e)Oblique random forest ensemble via Least Square Estimation for time series forecasting

f)Towards generating random forests via extremely randomized trees

g)Benchmarking Ensemble Classifiers with Novel Co-Trained Kernal Ridge Regression and Random Vector Functional Link Ensembles

In addition, there are some excellent review on ensemble learning, which is also related to the topic of the paper:

h) Ensemble methods in Machine learning

i)ensemble based classifiers

j)Ensemble classification and regression-recent developments, applications and future directions

Many thanks for these suggestions, which have been used to extend our review of the literature.