# What *is* the Best Classifier?
# What *is* the Best Regressor?

Ali Alizadeh Mansouri Kunal Singh Sidhant Gupta

## Abstract

*This report investigates two questions. First, for a given selection of data sets, can we say what is the 'best' classifier or the 'best' regressor in terms of good predictions? How much does the answer depend on the particular selection of data sets? How much does the answer depend on our computational constraints? We investigate these questions using data sets from the UCI repository. Second, we compare the interpretability of a decision tree classifier to that of a convolutional neural network. We compare the decision tree visualization to 'activation maximization', a technique to gain insight into the kinds of inputs that deep neural networks respond to.*

The abstract above was drafted for you. You can submit your project without modifying it, except that you must add 1–2 sentences to characterize the main idea of your novelty component. The rest of this template contains a suggested document structure, but you can change it as you see fit. Apart from the abstract, You should of course delete the text, formulas, figures, and tables that are currently included in this template. The report must be 4 pages of content, plus an additional page for references (if applicable). Fitting your report into 4 pages may be difficult, as is often the case when writing papers.

## 1. Introduction

Here you should write an introduction to how you approached the project, as if you were writing a short research paper. The introduction should be concise but provide an overview of your goals, your methodology and also provide brief mention of your novelty component. Even though some of the goals may be mentioned in the abstract, and some of the methodology is specified in the project guidelines, you should still try to write this report if it were a paper, and person reading had not seen the project guidelines. Describe what you did in your own words, however—copying and pasting is not OK.

The introduction is not the place for detailed descriptions of data preprocessing, training, hyperparameter search, testing, or success metrics. It is OK to mention some specifics if it helps clarify, but the full details should be explained later on, in the Methodology and Experiments sections. Likewise you do not need to review your conclusions here—there is a final section for that.

The introduction should be 1–1.5 columns in length.

## 2. Methodology & Experimental Results

Here you'll explain the general aspects of your methodology for determining which method is best in terms of prediction performance. In other words, here you can explain aspects that are common to both classification and regression.

### 2.1. Classification Experiments

Here you'll list the classification data sets (see report guidelines) and explain any methodological aspects (models evaluated, evaluation metrics used, *etc.*) that are specific to classification. Then you should describe your experimental results, and reference any relevant figures and/or tables.

### 2.2. Regression Experiments

Here you'll list the regression data sets (see report guidelines) and explain any methodological aspects (models evaluated, evaluation metrics used, *etc.*) that are specific to regression. Then you should describe your experimental results, and reference any relevant figures and/or tables.

### 2.3. Interpretability Experiments

Here you'll review what you did to process the CIFAR data and train your models. You should try to give some example of what you saw, in a figure—just enough to get an idea and to support your conclusions about interpretability. You'll then state your [hopefully collective] opinion on the interpretability of these models in particular and on interpretability in general.

## 3. Conclusions

Here you should summarize your thoughts on the questions asked in the abstract. For which questions can you

GROUP #29

offer a conclusion or at least a strong opinion? If a colleague of yours were about to download a dataset like the kind you studied here, what classifier and training procedure would you recommend he/she use? What regressor would you recommend, if any? What model(s) performed the 'worst' in your view? And was the result of your 'novelty component'? Show that you understand what your experimental results imply and do not imply.

## A. Detailed experimental results

*Optional section.* Here you can place supplementary plots and tables if they are needed to support your conclusions from the main report. You can include up to 2 extra pages of such material. They do not count towards your 4-page count. However, the instructor and TAs should not be obligated to read this section to understand your conclusions, it should only be used to provide 'supplementary' details. For example, as a full table of your performance results (algorithms × datasets) for classification and regression may does not fit within the 4-page limit, you can put such results here. If you do not feel including extra figures is necessary, that is OK, just delete this section.

## B. Overview of project code and data

*Optional section.* This is a guide written by you to help the course staff. Here you can make a few brief comments to the course staff about where they should start when looking at your project code, *e.g.* how to run your scripts and what data files contain the experimental results you used to draw your conclusions. If your project code already has such information in an obvious place, such as a 'README.md' then this section is not necessary.

## C. Examples of LaTeX

This section contains some examples of LaTeX to help you get started. (You should delete this section in the final report.) This is a reference to Table 1 and Table 2. This is a reference to Figure 1 and Figure 2. This is a citation [2] and this is multiple citations [2, 1]. This is *italics* and **bold** text. This is a formula $\sum_{i=1}^{N}(y_i - \hat{y}_i(\mathbf{x}))^2$ that is inline with the text ('text style') and this is a formula that is displayed separately ('display style'):

$$\sum_{i=1}^{N}(y_i - \hat{y}_i(\mathbf{x}))^2$$

These are formulas with an associated equation number

$$\mathbf{x} = \left[x_1, x_2, \ldots, x_N\right]^T \tag{1}$$

$$\boldsymbol{\phi} = \left[\phi_1, \phi_2, \ldots, \phi_M\right]^T \tag{2}$$

and we can now refer back to (1) or to (2) like so.

| Method | Ultra-Clustering | Random Jungles |
|---|---|---|
| Theirs | Works OK | All your base |
| Yours | Works better | are belong to us! |
| Ours | Works best! | I can haz publication? |

Table 1. This is the caption of a column-width table.

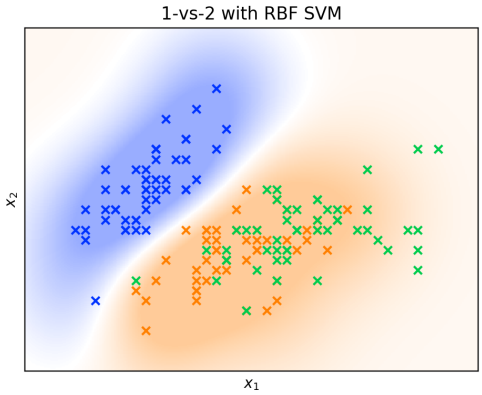


Figure 1. This is the caption of a column-width figure.

## References

[1] Christopher M Bishop. *Pattern recognition and machine learning*. springer, 2006 (cit. on p. 2).

[2] Leo Breiman et al. "Statistical modeling: The two cultures (with comments and a rejoinder by the author)". In: *Statistical science* 16.3 (2001), pp. 199–231 (cit. on p. 2).

| Method | Good? | Bad? | So-so? |
|---|---|---|---|
| Your method | Terrible | Yes, I made sure of it | Star Wars movies |
| My supervisor's old method (sigh) | I want Tim Horton's | People in hallway... | ...are talking too loudly |
| My proposed method | Yes, good! | No, I said good! | What? |

Table 2. This is the caption of a page-width table.

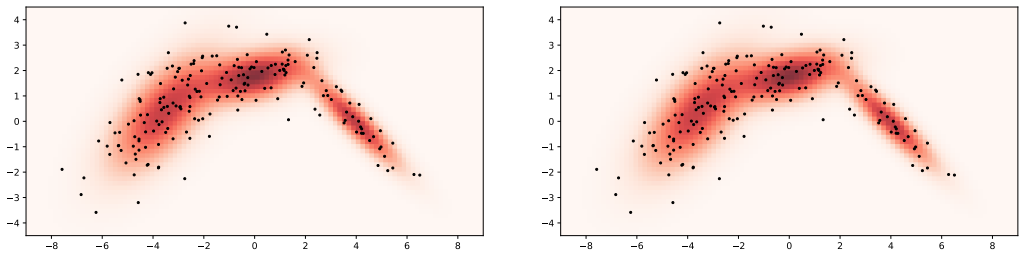

Figure 2. This is the caption of a page-width figure.