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. Moreover, we consider these questions for a non-stationary dataset that exhibits periodic concept drift. Third, 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.

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

As an extra step, we consider the Nebraska weather prediction data provided by the U.S. National Oceanic and Atmospheric Administration[1]. This a non-stationary dataset which exhibits periodic concept drift, meaning that the distribution of $p(y|X)$ changes over time. This poses certain challenges for all classical learning methods which assume that the distribution of data never changes, i.e., data is always stationary. Several research works on concept drift detection and adaptation, such as [1], have experimented on the dataset as a benchmark. We use the preprocessed version provided by the authors in [1][2]. The preprocessed version considers a binary classification problem (rain vs. no-rain). We compare and report the performance of the classification methods on the entire dataset.

2.2. 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 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.

[1] ftp://ftp.ncdc.noaa.gov/pub/data/gsod/

[2] http://users.rowan.edu/~polikar/nse.html

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.

References

[1] Ryan Elwell and Robi Polikar. "Incremental Learning of Concept Drift in Nonstationary Environments". In: *IEEE Transactions on Neural Networks* 22.10 (Oct. 2011), pp. 1517–1531. ISSN: 1045-9227, 1941-0093. DOI: 10/bgxfrz (cit. on p. 1).