# Evolving Fair Models: Fair and Accurate Machine Learning Models with Multi-objective Evolutionary Computing

**Julian F. Zucker** julian.zucker@gmail.com
Khoury College, Northeastern University, Boston, MA

**Clark Freifeld** c.freifeld@northeastern.edu
Khoury College, Northeastern University, Boston, MA

**Abstract**

Machine learning practitioners must ensure that their models are fair, a difficult task in the face of many conflicting definitions of fairness. We present a multi-objective optimization technique for training classification models that allows practitioners to identify the tradeoff space between accuracy and fairness. We explore four different fairness metrics. After optimizing for false positive rate, false negative rate, and one fairness metric, we generate Pareto frontiers of possible models with varying accuracy/fairness tradeoffs. This Pareto frontier can be used by machine learning practitioners to identify how much accuracy must be sacrificed to achieve a specific level of fairness. After optimizing for false positive rate, false negative rate, and two fairness metrics, we look into the tradeoffs between different fairness metrics. Understanding which fairness metrics are contradictory to or synonymous with others reduces the number of metrics which must be examined to understand the overall fairness of a model. The cost to accuracy of different levels of each fairness metric is computed, and the cost of greatly increasing fairness is shown to be small. r

**Keywords**

Genetic algorithms, evolutionary computing, data mining, binary classification, decision tree, machine learning fairness.

## 1 General Instructions

This document is a template which you can use to format your paper in preparation for publishing in the journal *Evolutionary Computation* (ECJ) and for submitting your paper to the journal. Our style file "ecj.sty" is compatible with LaTeX version 2e. Please note, also the first submission must be typeset using ECJ's LaTeX style. *Documents typeset by MS Word or other office programs are not accepted.*

Please make sure your paper is as complete and accurate as possible. You should typeset your paper as you would like to see it in its finally printed journal version (no double spacing; figures, tables should be embedded in the text at the most desirable locations).

Please provide author(s) first initial and last name(s) for the even-page running headline. Also provide a brief paper title (45 characters/spaces or less) for the odd-page

 

running headline. See the ecjHeader section at the top of this document for placement of these items.

The rest of the document provides a few examples of references and citations, how to set up figures, discusses common problems, and provides some general advice on writing your paper.

1. Give full names for authors. (T. Bones should be Tom Bones or Thomas, unless your first name is a military secret and everyone calls you "T".)
2. Be sure to provide 5 to 10 keywords for your paper. See "keywords" section above.
3. Use the *Evolutionary Computation* format for references. In text citations should use the authors names (Smith, 1997) or "Smith (1997) states ..." The references at the back of the paper should also follow the *Evolutionary Computation* format. Using a bitex file, with **natbib.sty** and **apalike.bst** is highly recommended. You should then use

   - `\cite{Smith} states that ...` to obtain "**?** states that ..."
   - `as stated in \citep{Smith}, ...` to obtain "as stated in (**?**) ..."

   If you don't use natbib, be sure to follow the rules:

   - 1 author: (Antonisse, 1989)
   - 2 authors: (Juliany and Vose, 1994)
   - multiple authors: (Reeves et al., 1990)
   - multiple citations: (Antonisse, 1989; Juliany and Vose, 1994)
4. If you use a bibtex file, please submit your bibtex file. In any case, please completely spell out journal and institution titles. Abbreviations may not be understood by all readers. Be sure to provide beginning *and* ending page numbers. Include editor(s) initials and last names(s), volume numbers, page ranges, publisher name and location.

   See the Reference section of this paper for specific examples.

## 2 About Figures

Figures potentially cause the most serious problems when processing LaTeX files. Image files should be in EPS (encapsulated postscript) or PNG format (both formats produce better outcome when enlarged). However, JPEG can be used as well. Make sure that no unusual files are required for processing your figures. The use of the "epsfig" or the related "graphicx" LaTeX package is strongly recommended as well as the format found in ecjsample.tex used for generating Figure 1 above. The use of the "includegraphics" command is commented out in the LaTeX file; but you can remove the "comment" symbols and use the commands shown in ecjsample.tex.

Make sure you provide an in-text reference to each of your figures. An example is Figure 1.

Make sure figures are clear and fit within the margins specified in the style file. Small figures with hard to read labels or hard to see lines are a common problem. Also

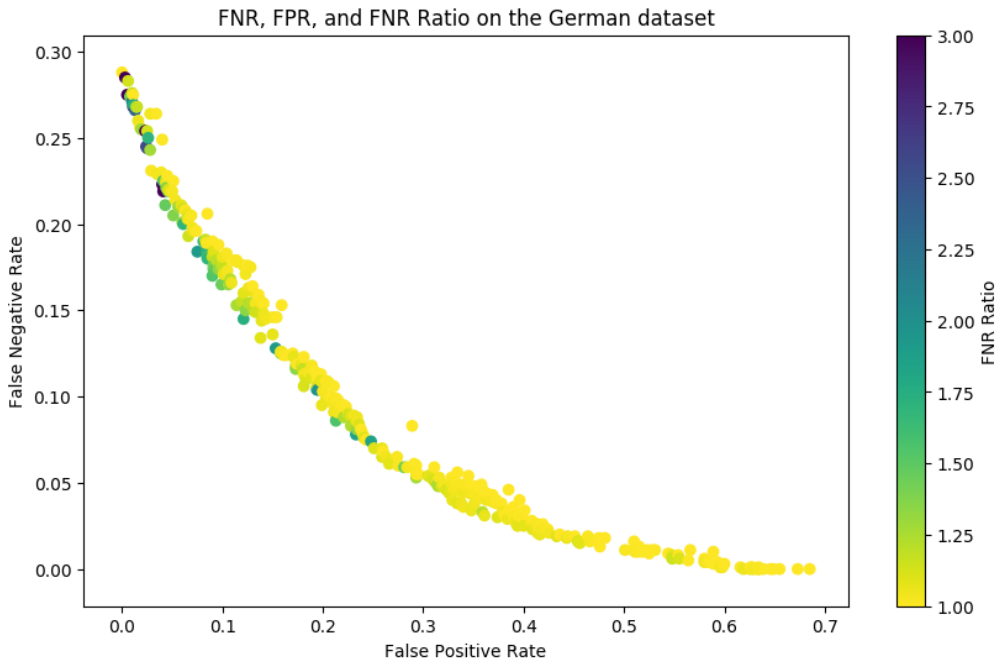


Figure 1: This is a caption below a framebox where a figure might appear.

make sure that labels and terms used in figures are defined in either the caption or the text. (It is best if they are defined in the caption.)

## 3 Common Problems and Advice

Avoid the use of too many acronyms. You may know the meaning and significance of BARF (e.g., Beer, Aspirin, Recreation, and Food), but this, in effect, amounts to the invention of a personal language that makes reading your paper more difficult. Define each acronym on its first occurrence it the text; thereafter, you may use the acronym alone.

Avoid run-on sentences. Typically these are very hard to parse. This is also one of the most common problems. While you are at it, use a spelling tool such as ispell.

Avoid paraphrases - "i.e.", "e.g.", "in other words...", or parenthetical information is often unnecessarily redundant.

Minimize the use of prepositions and adverbs to begin a sentence ("On the other hand...", "Conversely...", "Obviously..."). Vary your sentence structure.

Avoid using excessive supporting information/documentation. Select only the material that will strengthen your paper and is relevant and necessary.

For paper submission and the reviewing process, only the manuscript in PDF format is needed. After final acceptance, we will need all your files in electronic form.

This package should include at least the following files: .tex, .bib, .bbl, .eps, .pdf, and any special style files you have used to create your paper. We also need to be able to LaTeX and generate camera-ready PDF of your paper. Following the examples given in ecjsample.tex will help ensure your LaTeX file can be processed without error.

Thank you for submitting your work to *Evolutionary Computation*.

In Abellán and Castellano (2017) , binns Binns (2017) r

## References

Abellán, J. and Castellano, J. G. (2017). A comparative study on base classifiers in ensemble methods for credit scoring. *Expert Systems with Applications*, 73:1 – 10.

Binns, R. (2017). Fairness in machine learning: Lessons from political philosophy. *CoRR*, abs/1712.03586.