Race and the Machine

Re-examining Race and Ethnicity in Data Mining

Sophie Chou
MIT Media Lab
75 Amherst St
Cambridge, MA
soph@media.mit.edu

ABSTRACT

The rapid growth in popularity of machine learning and data mining to determine social outcomes – such as in healthcare, evidence-based sentencing, and credit assessment shows great potential, but presents many technical and ethical challenges. In particular, the analysis of race and ethnicity, either as a variable or by proxy, can be prone to profiling and discrimination. This paper aims to provoke new directions in approaching and re-approaching such data. Areas of challenge present themselves along the entire pipeline, including in the acquisition of biased datasets, feature and class selection for algorithmic classification and segmentation, and the application of results with accountability. Additionally, I discuss potential for computational solutions, both in monitoring existing software and detecting discrimination from datasets. I advocate for the treatment of race as a mutable social variable, and subsequent redesign of algorithms.

Categories and Subject Descriptors

H.4 [**Information Systems Applications**]: Miscellaneous; D.2.8 [**Software Engineering**]: Metrics—*complexity measures, performance measures*

General Terms

Theory

Keywords

data science, data mining, race, racism, discrimination

1. INTRODUCTION

Fill this out with stuff.

2. THE *BODY* OF THE PAPER

Typically, the body of a paper is organized into a hierarchical structure, with numbered or unnumbered headings for sections, subsections, sub-subsections, and even smaller

sections. The command `\section` that precedes this paragraph is part of such a hierarchy.[1] LaTeX handles the numbering and placement of these headings for you, when you use the appropriate heading commands around the titles of the headings. If you want a sub-subsection or smaller part to be unnumbered in your output, simply append an asterisk to the command name. Examples of both numbered and unnumbered headings will appear throughout the balance of this sample document.

Because the entire article is contained in the **document** environment, you can indicate the start of a new paragraph with a blank line in your input file; that is why this sentence forms a separate paragraph.

3. A BRIEF DEFINITION OF DATA MINING

With the growing availability of large datasets, there has been an increased interest in data-driven methods of analysis and prediction. The terms "data science", "data mining", "data analytics", and "machine learning" are often used interchangeably, and with confusion, to target similar problems.

To make a distinction, I will refer to data science as the broader umbrella concept, âĂIJthe set of activities involved in transforming collected data into valuable insights, products, or solutionsâĂİ [Cite DDS]; machine learning as the field of computer science that focuses on improving pattern recognition and prediction; data mining as the specific application of existing machine learning techniques to better understand and manipulate specific datasets, and data analytics as the business-facing aspects of the above.

In this paper, I choose to focus on data mining, for the key reason that it emphasizes choice of data and applications, which can have a great effect on potentially discriminatory results, even when using the same methods in machine learning to perform the analysis with. However, areas of challenge and redesign are relevant to all related fields.

4. THE POPULAR ACCOUNT OF RACE

Race, also a term that seems to expand and shrink in scope depending on context, is far trickier to define. Because it is deeply socially ingrained, race has the unfortunate quality of being viewed as intuitive, yet quickly reveals itself impossible

---

[1]This is the second footnote. It starts a series of three footnotes that add nothing informational, but just give an idea of how footnotes work and look. It is a wordy one, just so you see how a longish one plays out.

to define. Conceptions of race generally boil down into two camps: essentialist versus constructivist perspectives.

Common societal perceptions of race view it as a fixed and immutable variable. Although discussions have broadened in recent years, by and large it is considered a descriptive, rather than prescriptive term. A person is assigned a racial category before birth, chosen from a number of set possible classes, which remains constant throughout their entire life. This is known as an essentialist or primordialist notion of race, also deemed by philosopher Lawrence Blum as the âĂIJpopular accountâĂİ. Notably, it is distinct from other large-scale human groupings, such as ethnicity and nationality, which are viewed as cultivated (98). A main contributor to this notion of race is the fact that it includes physical or biological âĂIJmarkersâĂİ, traits that are prone to genetic explanation (although not necessarily so).

Social constructionists view race (along with other labels, such as gender) in a less deterministic way. Instead, it is seen as a category constructed and shaped by existing social structures. Race, therefore, is neither biologically nor genetically justified. It is contextually dependent on time, place, and social surrounding. Key to a constructionist view is the fact that race is not necessarily consistent.

5. RELATED WORK

In âĂIJRace as a âĂŸBundle of SticksâĂŹ: Designs that Estimate Effects of Seemingly Immutable CharacteristicsâĂİ, Harvard researchers Sen and Wasow address current limitations of studying race in the field of political science. Treating race as an immutable, pre-assigned trait limits experimental designâĂŤ since the variable is determined before birth, considering other effects post-birth risks introducing post-treatment bias, making studies of race difficult [Sen 2]. As a solution, Sen and Wasow suggest examining questions involving race in a more constructivist frameworkâĂŤ understanding it as a composite variable (âĂIJbundle of sticksâĂİ) so that each aspect may be examined separately in experiment for estimating casual inference [?].

Data mining faces similar issues in considering race as an immutable variable, challenges that compound when applying results to make predictions based on race. Although predictive algorithms are not necessarily concerned with determining causal links, they are concerned with inference at largeâĂŤ and more importantly, applications. In this paper, I expand on some of the issues specific to data mining in treating race as a static variable.

Big DataâĂŹs Disparate Impact, a survey by Barocas and Selbst, examines how existing practices in data mining are prone to discrimination, focusing on legal implications and difficulties with legislation [Cite]. Here, I choose to highlight how changes in algorithmic design can lead to positive effectâĂŤ focusing on the power of norms ahead of the law [Lessig].

Finally, there exists a small (but growing!) group of publications within the computational community criticizing popular techniques for demographic segmentation and inference. I highlight a few of these papers as examples for improvement.

6. VISION AND PROBLEM STATEMENT

In this paper I propose a shift in the data mining and machine learning community to treat race as a socially constructed variable, instead of one that is static and immutable. Existing methods of labeling race as such can be prone to stereotyping and profiling, and when applied to social applications, can have unwanted discriminatory effects. Reconsideration of the way race is examined in data, and how it is presented in results opens up new avenues for creativity without discounting existing techniques.

This consideration is both more experimentally responsible and scientifically sound. Regardless of personal philosophy, it is more appropriate to treat race in a constructionist sense when dealing with data mining, since it is, by definition, the mining of observed behaviors within a social context.

Accuracy suffers when behaviors that shift depending on social context are used as a proxy for variables seen as immutable and static. Moreover, harmful racial stereotyping and profiling may occur with the application of such algorithms. In the following, I will outline challenge areas in analyzing (and predicting) with racial data, suggesting examples contrary to the common practice that show improved results, and shine light on the potential for algorithmic and societal accountability through data mining.

6.1 Citations

Citations to articles [**?**, 2, 1, 3], conference proceedings [2] or books [5, 4] listed in the Bibliography section of your article will occur throughout the text of your article. You should use BibTeX to automatically produce this bibliography; you simply need to insert one of several citation commands with a key of the item cited in the proper location in the `.tex` file [4]. The key is a short reference you invent to uniquely identify each work; in this sample document, the key is the first author's surname and a word from the title. This identifying key is included with each item in the `.bib` file for your article.

The details of the construction of the `.bib` file are beyond the scope of this sample document, but more information can be found in the *Author's Guide*, and exhaustive details in the *LaTeX User's Guide*[4].

This article shows only the plainest form of the citation command, using `\cite`. This is what is stipulated in the SIGS style specifications. No other citation format is endorsed or supported.

6.2 Tables

Because tables cannot be split across pages, the best placement for them is typically the top of the page nearest their initial cite. To ensure this proper "floating" placement of tables, use the environment **table** to enclose the table's contents and the table caption. The contents of the table itself must go in the **tabular** environment, to be aligned properly in rows and columns, with the desired horizontal and vertical rules. Again, detailed instructions on **tabular** material is found in the *LaTeX User's Guide*.

Immediately following this sentence is the point at which Table 1 is included in the input file; compare the placement of the table here with the table in the printed dvi output of this document.

To set a wider table, which takes up the whole width of the page's live area, use the environment **table*** to enclose the table's contents and the table caption. As with a single-column table, this wide table will "float" to a location deemed more desirable. Immediately following this sentence

Table 1: Frequency of Special Characters

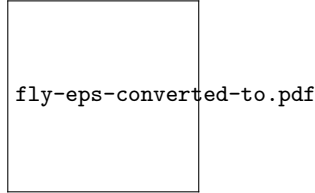| Non-English or Math	Frequency	Comments
Ø	1 in 1,000	For Swedish names
π	1 in 5	Common in math
$	4 in 5	Used in business
Ψ_1^2	1 in 40,000	Unexplained usage

fly-eps-converted-to.pdf

Figure 1: A sample black and white graphic (.eps format).

is the point at which Table 2 is included in the input file; again, it is instructive to compare the placement of the table here with the table in the printed dvi output of this document.

6.3 Figures

Like tables, figures cannot be split across pages; the best placement for them is typically the top or the bottom of the page nearest their initial cite. To ensure this proper "floating" placement of figures, use the environment **figure** to enclose the figure and its caption.

This sample document contains examples of **.eps** and **.ps** files to be displayable with LaTeX. More details on each of these is found in the *Author's Guide*.

As was the case with tables, you may want a figure that spans two columns. To do this, and still to ensure proper "floating" placement of tables, use the environment **figure*** to enclose the figure and its caption. and don't forget to end the environment with figure*, not figure!

Note that either **.ps** or **.eps** formats are used; use the \epsfig or \psfig commands as appropriate for the different file types.

6.4 Theorem-like Constructs

Other common constructs that may occur in your article are the forms for logical constructs like theorems, axioms, corollaries and proofs. There are two forms, one produced by the command \newtheorem and the other by the command \newdef; perhaps the clearest and easiest way to distinguish them is to compare the two in the output of this sample document:
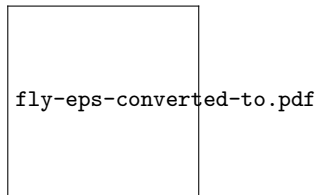

fly-eps-converted-to.pdf

Figure 2: A sample black and white graphic (.eps format) that has been resized with the epsfig command.

Figure 4: A sample black and white graphic (.ps format) that has been resized with the psfig command.

This uses the **theorem** environment, created by the \newtheorem command:

THEOREM 1. *Let f be continuous on* $[a, b]$*. If G is an antiderivative for f on* $[a, b]$*, then*

$$\int_a^b f(t)dt = G(b) - G(a).$$

The other uses the **definition** environment, created by the \newdef command:

Definition 1. If z is irrational, then by e^z we mean the unique number which has logarithm z:

$$\log e^z = z$$

Two lists of constructs that use one of these forms is given in the *Author's Guidelines.*

There is one other similar construct environment, which is already set up for you; i.e. you must *not* use a \newdef command to create it: the **proof** environment. Here is a example of its use:

PROOF. Suppose on the contrary there exists a real number L such that

$$\lim_{x \to \infty} \frac{f(x)}{g(x)} = L.$$

Then

$$l = \lim_{x \to c} f(x) = \lim_{x \to c} \left[gx \cdot \frac{f(x)}{g(x)} \right] = \lim_{x \to c} g(x) \cdot \lim_{x \to c} \frac{f(x)}{g(x)} = 0 \cdot L = 0,$$

which contradicts our assumption that $l \neq 0$. □

Complete rules about using these environments and using the two different creation commands are in the *Author's Guide*; please consult it for more detailed instructions. If you need to use another construct, not listed therein, which you want to have the same formatting as the Theorem or the Definition[5] shown above, use the \newtheorem or the \newdef command, respectively, to create it.

A *Caveat* for the TeX Expert

Because you have just been given permission to use the \newdef command to create a new form, you might think you can use TeX's \def to create a new command: *Please refrain from doing this!* Remember that your LaTeX source code is primarily intended to create camera-ready copy, but may be converted to other forms – e.g. HTML. If you inadvertently omit some or all of the \defs recompilation will be, to say the least, problematic.

7. CONCLUSIONS

This paragraph will end the body of this sample document. Remember that you might still have Acknowledgments or Appendices; brief samples of these follow. There is still the Bibliography to deal with; and we will make a disclaimer about that here: with the exception of the reference to the LaTeX book, the citations in this paper are to articles which have nothing to do with the present subject and are used as examples only.

Table 2: Some Typical Commands

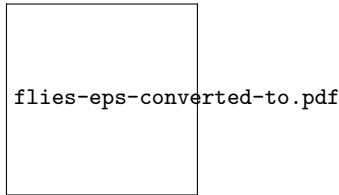| Command	A Number	Comments
`\alignauthor`	100	Author alignment
`\numberofauthors`	200	Author enumeration
`\table`	300	For tables
`\table*`	400	For wider tables

```
flies-eps-converted-to.pdf
```

Figure 3: A sample black and white graphic (.eps format) that needs to span two columns of text.

8. ACKNOWLEDGMENTS

9. REFERENCES

[1] J. Braams. Babel, a multilingual style-option system for use with latex's standard document styles. *TUGboat*, 12(2):291–301, June 1991.

[2] M. Clark. Post congress tristesse. In *TeX90 Conference Proceedings*, pages 84–89. TeX Users Group, March 1991.

[3] M. Herlihy. A methodology for implementing highly concurrent data objects. *ACM Trans. Program. Lang. Syst.*, 15(5):745–770, November 1993.

[4] L. Lamport. *LaTeX User's Guide and Document Reference Manual.* Addison-Wesley Publishing Company, Reading, Massachusetts, 1986.

[5] S. Salas and E. Hille. *Calculus: One and Several Variable.* John Wiley and Sons, New York, 1978.

APPENDIX

A. HEADINGS IN APPENDICES

The rules about hierarchical headings discussed above for the body of the article are different in the appendices. In the **appendix** environment, the command **section** is used to indicate the start of each Appendix, with alphabetic order designation (i.e. the first is A, the second B, etc.) and a title (if you include one). So, if you need hierarchical structure *within* an Appendix, start with **subsection** as the highest level. Here is an outline of the body of this document in Appendix-appropriate form:

A.1 Introduction

A.2 The Body of the Paper

A.2.1 Type Changes and Special Characters

A.2.2 Math Equations

Inline (In-text) Equations.

Display Equations.

A.2.3 Citations

A.2.4 Tables

A.2.5 Figures

A.2.6 Theorem-like Constructs

A Caveat for the TeX Expert

A.3 Conclusions

A.4 Acknowledgments

A.5 Additional Authors

This section is inserted by LaTeX; you do not insert it. You just add the names and information in the `\additionalauthors` command at the start of the document.

A.6 References

Generated by bibtex from your .bib file. Run latex, then bibtex, then latex twice (to resolve references) to create the .bbl file. Insert that .bbl file into the .tex source file and comment out the command `\thebibliography`.

B. MORE HELP FOR THE HARDY

The sig-alternate.cls file itself is chock-full of succinct and helpful comments. If you consider yourself a moderately experienced to expert user of LaTeX, you may find reading it useful but please remember not to change it.