

Explaining black-box algorithms using CounterfactualExplanations.jl

Patrick Altmeyer¹ and Cynthia Liem¹

¹Delft University of Technology

ABSTRACT

Machine learning models like deep neural networks have become so complex and opaque over recent years that they are generally considered as black boxes. Nonetheless such models play a key role in modern automated decision-making systems. Counterfactual explanations (CE) can help programmers make sense of the systems they build: they explain how inputs into a system need to change for it to produce different decisions. Explanations that involve realistic and actionable changes can be used for the purpose of algorithmic recourse (AR): they offer individuals subject to algorithms a way to turn a negative decision into positive one. In this article we discuss the usefulness of counterfactual explanations for interpretable machine learning and demonstrate its implementation in Julia using the CounterfactualExplanations package.

Keywords

Julia, Interpretable Machine Learning, Counterfactual Explanations, Algorithmic Recourse

1. Introduction

2. Methodology

The juliacon class file preserves the standard LATEX interface such that any document that can be produced using the standard LATEX article class can also be produced with the class file.

It is likely that the make up will change after file submission. For this reason, we ask you to ignore details such as slightly long lines, page stretching, or figures falling out of synchronization, as these details can be dealt with at a later stage.

Use should be made of symbolic references (`\ref`) in order to protect against late changes of order, etc.

3. Using CounterfactualExplanations

If the file `juliacon.cls` is not already in the appropriate system directory for L^AT_EX files, either arrange for it to be put there or copy it to your working directory. The `juliacon` document class is implemented as a complete class, not a document style option. In order to use the `juliacon` document class, replace `\verbarticle` by `juliacon` in the `\documentclass` command at the beginning of your document:

```
\documentclass{article}
```

replace by

```
\documentclass{juliacon}
```

In general, the following standard document style options should *not* be used with the *article* class file:

- (1) `10pt`, `11pt`, `12pt` ? unavailable;
- (2) `twoside` (no associated style file) ? `twoside` is the default;
- (3) `fleqn`, `leqno`, `titlepage` ? should not be used;

4. Additional Document Style Options

The following additional style option is available with the `juliacon` class file:

Please place any additional command definitions at the very start of the L^AT_EX file, before the `\begin{document}`. For example, user-defined `\def` and `\newcommand` commands that define macros for technical expressions should be placed here. Other author-defined macros should be kept to a minimum.

Commands that differ from the standard L^AT_EX interface, or that are provided in addition to the standard interface, are explained in this guide. This guide is not a substitute for the L^AT_EX manual itself. Authors planning to submit their papers in L^AT_EX are advised to use `juliacon.cls` as early as possible in the creation of their files.

5. Additional features

In addition to all the standard L^AT_EX design elements, the `juliacon` class file includes the following features: In general, once you have used the additional `juliacon.cls` facilities in your document, do not process it with a standard L^AT_EX class file.

5.1 Titles, Author's Name, and Affiliation

The title of the article, author's name, and affiliation are used at the beginning of the article (for the main title). These can be produced using the following code:

```
\title{ This is an example of article title} }
\author{
  \large 1st Author \\[ -3pt]
  \normalsize 1st author's affiliation \\[ -3pt]
  \normalsize 1st line of address \\[ -3pt]
  \normalsize 2nd line of address \\[ -3pt]
  \normalsize 1st author's email address \\[ -3pt]
  \and
  \large 2nd Author \\[ -3pt]
  \normalsize 2nd author's affiliation \\[ -3pt]
```

Table 1. If necessary, the tables can be extended both columns.

Label	Description	Number of Users	Number of Queries
Test 1	Training Data	70	104
Test 2	Testing Data I		105
Test 3	Testing Data II	30	119
	Total	100	328

This is an example of table footnote.

```

\normalsize 1st line of address \[-3pt]
\normalsize 2nd line of address \[-3pt]
\normalsize 2nd author's email address \[-3pt]
\and
\large 3rd Author \[-3pt]
\normalsize 3rd author's affiliation \[-3pt]
\normalsize 1st line of address \[-3pt]
\normalsize 2nd line of address \[-3pt]
\normalsize 3rd author's email address \[-3pt]
}
\maketitle

```

5.2 Writing Julia code

A special environment is already defined for Julia code, built on top of *listings* and *jlcode*.

```

\begin{lstlisting}[language = Julia]
using Plots

x = -3.0:0.01:3.0
y = rand(length(x))
plot(x, y)
\end{lstlisting}

```

```

using Plots

x = -3.0:0.01:3.0
y = rand(length(x))
plot(x, y)

```

5.3 Abstracts, Key words, term etc...

At the beginning of your article, the title should be generated in the usual way using the `\maketitle` command. For general term and keywords use `\terms`, `\keywords` commands respectively. The abstract should be enclosed within an abstract environment. All these environment can be produced using the following code:

```

\terms{Experimentation, Human Factors}

\keywords{Face animation, image-based modelling...}

\begin{abstract}
In this paper, we propose a new method for the
systematic determination of the model's base of
time varying delay system. This method based on
the construction of the classification data related
to the considered system. The number, the orders,
the time delay and the parameters of the local
models are generated automatically without any
knowledge about the full operating range of the
process. The parametric identification of the local
models is realized by a new recursive algorithm for

```

on line identification of systems with unknown time delay. The proposed algorithm allows simultaneous estimation of time delay and parameters of discrete-time systems. The effectiveness of the new method has been illustrated through simulation.

6. Some guidelines

The following notes may help you achieve the best effects with the `juliacon` class file.

6.1 Sections

$\LaTeX 2_{\epsilon}$ provides four levels of section headings and they are all defined in the `juliacon` class file:

```

—\section
—\subsection
—\subsubsection
—\paragraph

```

Section headings are automatically converted to allcaps style.

6.2 Lists

The `juliacon` class file provides unnumbered lists using the `unnumlist` environment for example,

First unnumbered item which has no label and is indented from the left margin.

Second unnumbered item.

Third unnumbered item.

The unnumbered list which has no label and is indented from the left margin. was produced by:

```

\begin{unnumlist}
\item First unnumbered item...
\item Second unnumbered item...
\item Third unnumbered item...
\end{unnumlist}

```

The `juliacon` class file also provides hyphen list using the `itemize` environment for example,

```

—First unnumbered bulleted item which has no label and is
  indented from the left margin.
—Second unnumbered bulleted item.
—Third unnumbered bulleted item which has no label and is
  indented from the left margin.

```

was produced by:



Fig. 1. This is example of the image in a column.

```
\begin{itemize}
\item First item...
\item Second item...
\item Third item...
\end{itemize}
```

Numbered list is also provided in acmtog class file using the enumerate environment for example,

- (1) The attenuated and diluted stellar radiation.
- (2) Scattered radiation, and
- (3) Reradiation from other grains.

was produced by:

```
\begin{enumerate}
\item The attenuated...
\item Scattered radiation, and...
\item Reradiation from other grains...
\end{enumerate}
```

6.3 Illustrations (or figures)

The juliacon class file will cope with most of the positioning of your illustrations and you should not normally use the optional positional qualifiers on the figure environment that would override these decisions.

The figure 1 is taken from the JuliaGraphs organization ¹.

Figure captions should be *below* the figure itself, therefore the \caption command should appear after the figure or space left for an illustration. For example, Figure 1 is produced using the following commands:

```
\begin{figure}
\centerline{\includegraphics[width=20pc]{Graphics.eps}}
\caption{An example of the testing process for a
binary tree. The global null hypothesis is tested
first at level  $\alpha$  (a), and the level of
individual variables is reached last (d). Note
that individual hypotheses can be tested at
level  $\alpha/4$  and not  $\alpha/8$  as one might
expect at first.}
\label{sample-figure_2}
\end{figure}
```

Figures can be resized using first and second argument of \includegraphics command. First argument is used for modifying figure height and the second argument is used for modifying figure width respectively.

¹<https://github.com/JuliaGraphs>

Table 2. Tuning Set and Testing Set

Label	Description	Number of Users	Number of Queries
Test 1	Training Data	70	104
Test 2	Testing Data I		105
Test 3	Testing Data II	30	119
	Total	100	328

Cross-referencing of figures, tables, and numbered, displayed equations using the \label and \ref commands is encouraged. For example, in referencing Figure 1 above, we used `Figure~\ref{sample-figure}`

6.4 Tables

The juliacon class file will cope with most of the positioning of your tables and you should not normally use the optional positional qualifiers on the table environment which would override these decisions. Table captions should be at the top.

```
\begin{table}
\tbl{Tuning Set and Testing Set}{
\begin{tabular}{|l|l|c|c|}\hline
Label & \multicolumn{1}{c|}{Description}
& Number of Users &
Number of Queries\\\hline
Train70 & Training Data &
\smash{\raise-7pt\hbox{70}} & 104\\
\cline{1-2}\cline{4-4}
Test70 & Testing Data I & & 105\\\hline
Test30 & Testing Data II & 30 & 119\\\hline
& Total & 100 & 328\\\hline
\end{tabular}}
\end{table}
```

6.5 Landscaping Pages

If a table is too wide to fit the standard measure, it may be turned, with its caption, to 90 degrees. Landscape tables cannot be produced directly using the juliacon class file because T_EX itself cannot turn the page, and not all device drivers provide such a facility. The following procedure can be used to produce such pages.

Use the package rotating in your document and change the coding from

```
\begin{table}...\end{table}
to
\begin{sidewaystable}...\end{sidewaystable}
and for figures
\begin{figure}...\end{figure}
to
\begin{sidewaysfigure}...\end{sidewaysfigure}
```

environments in your document to turn your table on the appropriate page of your document. For instance, the following code prints a page with the running head, a message half way down and the table number towards the bottom.

```
\begin{sidewaystable}
\tbl{Landscape table caption to go here.}{...}
\label{landtab}
\end{sidewaystable}
```

6.6 Double Column Figure and Tables

For generating the output of figures and tables in double column we can use the following coding:

(1) For Figures:

```
\begin{figure*}...\end{figure*}
```

(2) For landscape figures:

```
\begin{sidewaysfigure*}...\end{sidewaysfigure*}
```

(3) For Tables:

```
\begin{table*}...\end{table*}
```

(4) For landscape tables:

```
\begin{sidewaystable*}...\end{sidewaystable*}
```

6.7 Typesetting Mathematics

The juliacon class file will set displayed mathematics with center to the column width, provided that you use the $\text{\LaTeX} 2_{\epsilon}$ standard of open and closed square brackets as delimiters. The equation

$$\sum_{i=1}^p \lambda_i = (S)$$

was typeset using the acmtog class file with the commands

```
\[
\sum_{i=1}^p \lambda_i = (S)
\]
```

For display equations, cross-referencing is encouraged. For example,

```
\begin{equation}
(n-1)^{-1} \sum_{i=1}^n (X_i - \overline{X})^2.
\label{eq:samplevar}
\end{equation}
Equation~(\ref{eq:samplevar}) gives the formula for
sample variance.
```

The following output is generated with the above coding:

$$(n-1)^{-1} \sum_{i=1}^n (X_i - \overline{X})^2. \quad (1)$$

Equation (1) gives the formula for sample variance.

6.8 Enunciations

The juliacon class file generates the enunciations with the help of the following commands:

```
\begin{theorem}...\end{theorem}
\begin{strategy}...\end{strategy}
\begin{property}...\end{property}
\begin{proposition}...\end{proposition}
\begin{lemma}...\end{lemma}
\begin{example}...\end{example}
\begin{proof}...\end{proof}
\begin{definition}...\end{definition}
\begin{algorithm}...\end{algorithm}
\begin{remark}...\end{remark}
```

The above-mentioned coding can also include optional arguments such as

```
\begin{theorem}[...]. Example for theorem:
\begin{theorem}[Generalized Poincare Conjecture]
Four score and seven ... created equal.
\end{theorem}
```

THEOREM 1 GENERALIZED POINCARÉ CONJECTURE. *Four score and seven years ago our fathers brought forth, upon this continent, a new nation, conceived in Liberty, and dedicated to the proposition that all men are created equal.*

6.9 Extract

Extract environment should be coded within

```
\begin{extract}...\end{extract}
```

6.10 Balancing column at last page

For balancing the both column length at last page use :

```
\vadjust{\vfill\pagebreak}
```

at appropriate place in your \TeX file or in bibliography file.

7. Handling references

References are most easily (and correctly) generated using the BIB-TEX, which is easily invoked via

```
\bibliographystyle{juliacon}
\bibliography{ref}
```

When submitting the document source (.tex) file to external parties, the ref.bib file should be sent with it. [?]

8. References