PROJECT REPORT (HINF6210-DATAMINING)

Predicting the ICU admissions

Edwin Kagereki

^{a}Dalhousie University

ABSTRACT
This template is for authors who are preparing a manuscript for a Taylor & Francis journal using the LaTeX document preparation system and the 'interact} class file, which is available via selected journals' home pages on the Taylor & Francis website.

KEYWORDS
Sections; lists; figures; tables; mathematics; fonts; references; appendices

1. Introduction

2. Section II

3. Project dataset, Code, and Sample outputs.

3.1. *Dataset*

MIMIC-III is a large, freely-available database comprising deidentified health-related data associated with over forty thousand patients who stayed in critical care units of the Beth Israel Deaconess Medical Center between 2001 and 2012. The database includes information such as demographics, vital sign measurements made at the bedside (~1 data point per hour), laboratory test results, procedures, medications, caregiver notes, imaging reports, and mortality (including post-hospital discharge).

MIMIC supports a diverse range of analytic studies spanning epidemiology, clinical decision-rule improvement, and electronic tool development. It is notable for three factors: it is freely available to researchers worldwide; it encompasses a diverse and very large population of ICU patients; and it contains highly granular data, including vital signs, laboratory results, and medications.

This report is organized as follows. In section II, the report describes the project statement and objectives. In section III the report describes the solution to the problem including the dataset, code and the outputs. In Section IV, the analysis report is presented A short summary concludes the report.

4. Project report

After completing the coding section find the analysis.

CONTACT Edwin Kagereki. Email: `kagereki@dal.ca`

4.1. *Introduction*

4.2. *Results*

4.2.1. Exploratory data analysis

MIMIC-III contains data associated with 58976 distinct ICU admissions for 50799 adult patients (aged 16 years or above), admitted to critical care units between 2001 and 2012. In addition, it contains data for 8110 neonates admitted between 2001 and 2008. The data covers 38,597 distinct adult patients and 49,785 hospital admissions. The median age of adult patients is 65.8 years (Q1–Q3: 52.8–77.8), 55.9% patients are male, and in-hospital mortality is 11.5%. The median length of an ICU stay is 2.1 days (Q1–Q3: 1.2–4.6) and the median length of a hospital stay is 6.9 days (Q1-Q3: 4.1–11.9). A mean of 4579 charted observations ('chartevents') and 380 laboratory measurements ('labevents') are available for each hospital admission. Table 1 provides a breakdown of the adult population by care unit.

```
## function (x, ...)
## {
##     UseMethod("table1")
## }
## <bytecode: 0x000000002dc83fc8>
## <environment: namespace:table1>
```

4.3. *Feauture selection*

Manuscripts for possible publication should be submitted to the Editors for review as directed in the journal's Instructions for Authors, and in accordance with any technical instructions provided in the journal's ScholarOne Manuscripts or Editorial Manager site. Your LaTeX source file(s), the class file and any graphics files will be required in addition to the final PDF version when final, revised versions of accepted manuscripts are submitted.

Please ensure that any author-defined macros used in your article are gathered together in the preamble of your .tex file, i.e. before the `\begin{document}` command. Note that if serious problems are encountered in the coding of a document (missing author-defined macros, for example), the typesetter may resort to rekeying it.

4.4. *Feature engineering*

5. Using the `interact` class file

For convenience, simply copy the `interact.cls` file into the same directory as your manuscript files (you do not need to install it in your TeX distribution). In order to use the `interact` document class, replace the command `\documentclass{article}` at the beginning of your document with the command `\documentclass{interact}`.

The following document-class options should *not* be used with the `interact` class file:

- `10pt`, `11pt`, `12pt` – unavailable;
- `oneside`, `twoside` – not necessary, `oneside` is the default;
- `leqno`, `titlepage` – should not be used;

- `twocolumn` – should not be used (see Subsection~**??**);
- `onecolumn` – not necessary as it is the default style.

To prepare a manuscript for a journal that is printed in A4 (two column) format, use the `largeformat` document-class option provided by `interact.cls`; otherwise the class file produces pages sized for B5 (single column) format by default. The `geometry` package should not be used to make any further adjustments to the page dimensions.

5.1. *Model selection*

6. Additional features of the `interact` class file

6.1. *Title, authors' names and affiliations, abstracts and article types*

6.2. *Abbreviations*

A list of abbreviations may be included if required, enclosed within an `abbreviations` environment, i.e. `\begin{abbreviations}`... `\end{abbreviations}`, immediately following the `abstract` environment.

6.3. *Keywords*

A list of keywords may be included if required, enclosed within a `keywords` environment, i.e. `\begin{keywords}`... `\end{keywords}`. Additional keywords in other languages (preceded by a translation of the word `keywords`) may also be included within the `keywords` environment if required.

6.4. *Subject classification codes*

AMS, JEL or PACS classification codes may be included if required. The `interact` class file provides an `amscode` environment, i.e. `\begin{amscode}`... `\end{amscode}`, a `jelcode` environment, i.e. `\begin{jelcode}`... `\end{jelcode}`, and a `pacscode` environment, i.e. `\begin{pacscode}`... `\end{pacscode}` to assist with this.

6.5. *Additional footnotes to the title or authors' names*

Note that any `footnotes` to the main text will automatically be assigned the superscript symbols 1, 2, 3, etc. by the class file.[1]

7. Embedding R code

7.1. *Code chunks*

```
summary(cars)
```

```
##      speed           dist
```

---

[1] If preferred, the `endnotes` package may be used to set the notes at the end of your text, before the bibliography. The symbols will be changed to match the style of the journal if necessary by the typesetter.

```
##  Min.   : 4.0   Min.   :  2.00
##  1st Qu.:12.0   1st Qu.: 26.00
##  Median :15.0   Median : 36.00
##  Mean   :15.4   Mean   : 42.98
##  3rd Qu.:19.0   3rd Qu.: 56.00
##  Max.   :25.0   Max.   :120.00
```

7.2. *Including Plots*

You can also embed plots, for example:
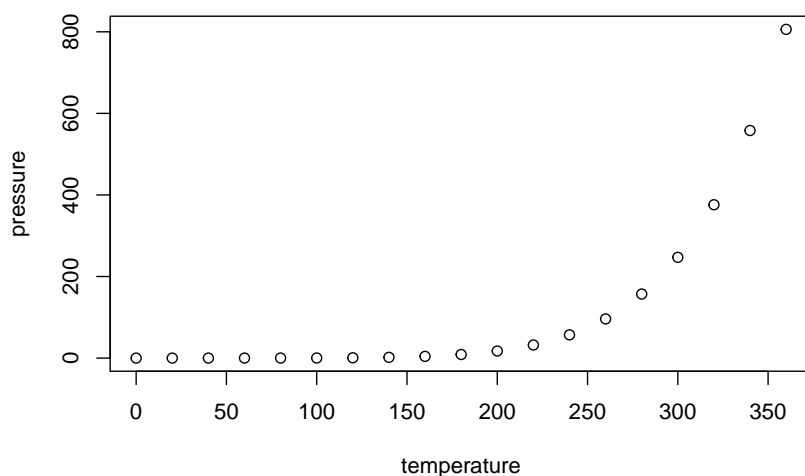
Note that the `echo = FALSE` parameter was added to the code chunk to prevent printing of the R code that generated the plot.

## 8. Some guidelines for using the standard features of LaTeX

8.1. *Sections*

The Interact layout style allows for five levels of section heading, all of which are provided in the `interact` class file using the standard LaTeX commands `\section`, `\subsection`, `\subsubsection`, `\paragraph` and `\subparagraph`. Numbering will be automatically generated for all these headings by default.

8.2. *Lists*

Numbered lists are produced using the `enumerate` environment, which will number each list item with arabic numerals by default. For example,

1. first item
2. second item
3. third item

Alternative numbering styles can be achieved by inserting an optional argument in

4

(a) An example of an individual figure sub-caption.

(b) A slightly shorter sub-caption.
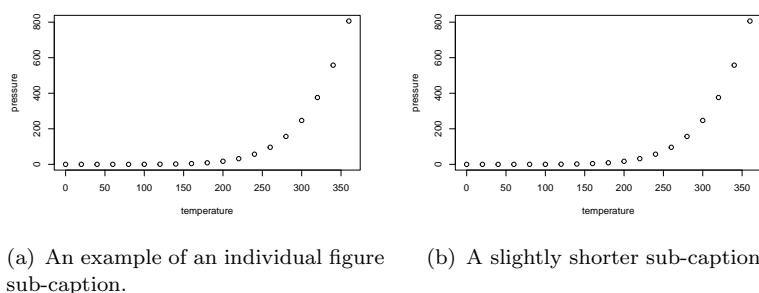
Figure 1. Example of a two-part figure with individual sub-captions showing that captions are flush left and justified if greater than one line of text.

square brackets to each `item`, e.g. `\item[(i)] first item`, to create a list numbered with roman numerals at level one.

Bulleted lists are produced using the `itemize` environment. For example,

- First bulleted item
- Second bulleted item
- Third bulleted item

8.3. *Figures*

```
plot(pressure)
```

The `interact` class file will deal with positioning your figures in the same way as standard LaTeX. It should not normally be necessary to use the optional `[htb]` location specifiers of the `figure` environment in your manuscript; you may, however, find the `[p]` placement option or the `endfloat` package useful if a journal insists on the need to separate figures from the text.

Figure captions appear below the figures themselves, therefore the `\caption` command should appear after the body of the figure. For example, Figure~1 with caption and sub-captions is produced using the following commands:

```
\begin{figure}
\centering
\subfigure[An example of an individual figure sub-caption.]{%
\resizebox*{5cm}{!}{\includegraphics{path/to/fig}}}\hspace{5pt}
\subfigure[A slightly shorter sub-caption.]{%
\resizebox*{5cm}{!}{\includegraphics{path/to/fig}}}
\caption{Example of a two-part figure with individual sub-captions
 showing that captions are flush left and justified if greater
 than one line of text.} \label{sample-figure}
\end{figure}
```

To ensure that figures are correctly numbered automatically, the `\label` command should be included just after the `\caption` command, or in its argument.

The `\subfigure` command requires `subfigure.sty`, which is called in the preamble of the `interacttfssample.tex` file (to allow your choice of an alternative package if

5

Table 1. Example of a table showing that its caption is as wide as the table itself and justified.

Class	Type					
	One	Two	Three	Four	Five	Six
Alpha[a]	A1	A2	A3	A4	A5	A6
Beta	B2	B2	B3	B4	B5	B6
Gamma	C2	C2	C3	C4	C5	C6

[a]This footnote shows how to include footnotes to a table if required.

preferred) and included in the Interact LaTeX bundle for convenience. Please supply any additional figure macros used with your article in the preamble of your .tex file.

The source files of any figures will be required when the final, revised version of a manuscript is submitted. Authors should ensure that these are suitable (in terms of lettering size, etc.) for the reductions they envisage.

The `epstopdf` package can be used to incorporate encapsulated PostScript (.eps) illustrations when using PDFLaTeX, etc. Please provide the original .eps source files rather than the generated PDF images of those illustrations for production purposes.

8.4. *Tables*

The `interact` class file will deal with positioning your tables in the same way as standard LaTeX. It should not normally be necessary to use the optional `[htb]` location specifiers of the `table` environment in your manuscript; you may, however, find the `[p]` placement option or the `endfloat` package useful if a journal insists on the need to separate tables from the text.

The `tabular` environment can be used as shown to create tables with single horizontal rules at the head, foot and elsewhere as appropriate. The captions appear above the tables in the Interact style, therefore the `\tbl` command should be used before the body of the table. For example, Table~1 is produced using the following commands:

```
\begin{table}
\tbl{Example of a table showing that its caption is as wide as
 the table itself and justified.}
{\begin{tabular}{lcccccc} \toprule
 & \multicolumn{2}{l}{Type} \\ \cmidrule{2-7}
 Class & One & Two & Three & Four & Five & Six \\ \midrule
 Alpha\textsuperscript{a} & A1 & A2 & A3 & A4 & A5 & A6 \\
 Beta & B2 & B2 & B3 & B4 & B5 & B6 \\
 Gamma & C2 & C2 & C3 & C4 & C5 & C6 \\ \bottomrule
\end{tabular}}
\tabnote{\textsuperscript{a}This footnote shows how to include
 footnotes to a table if required.}
\label{sample-table}
\end{table}
```

To ensure that tables are correctly numbered automatically, the `\label` command should be included just before `\end{table}`.

The `\toprule`, `\midrule`, `\bottomrule` and `\cmidrule` commands are those used by `booktabs.sty`, which is called by the `interact` class file and included in the Interact

LaTeX bundle for convenience. Tables produced using the standard commands of the `tabular` environment are also compatible with the `interact` class file.

8.5.   *Landscape pages*

If a figure or table is too wide to fit the page it will need to be rotated, along with its caption, through 90° anticlockwise. Landscape figures and tables can be produced using the `rotating` package, which is called by the `interact` class file. The following commands (for example) can be used to produce such pages.

```
\setcounter{figure}{1}
\begin{sidewaysfigure}
\centerline{\epsfbox{figname.eps}}
\caption{Example landscape figure caption.}
\label{landfig}
\end{sidewaysfigure}

\setcounter{table}{1}
\begin{sidewaystable}
 \tbl{Example landscape table caption.}
  {\begin{tabular}{@{}llllcll}
    .
    .
    .
  \end{tabular}}\label{landtab}
\end{sidewaystable}
```

Before any such float environment, use the `\setcounter` command as above to fix the numbering of the caption (the value of the counter being the number given to the preceding figure or table). Subsequent captions will then be automatically renumbered accordingly. The `\epsfbox` command requires `epsfig.sty`, which is called by the `interact` class file and is also included in the Interact LaTeX bundle for convenience.

Please note that if the `endfloat` package is used, one or both of the commands

```
\DeclareDelayedFloatFlavor{sidewaysfigure}{figure}
\DeclareDelayedFloatFlavor{sidewaystable}{table}
```

will need to be included in the preamble of your .tex file, after the `endfloat` package is loaded, in order to process any landscape figures and/or tables correctly.

8.6.   *Theorem-like structures*

A predefined `proof` environment is provided by the `amsthm` package (which is called by the `interact` class file), as follows:

Proof. More recent algorithms for solving the semidefinite programming relaxation are particularly efficient, because they explore the structure of the MAX-CUT problem.
□

This was produced by simply typing:

```
\begin{proof}
```

```
More recent algorithms for solving the semidefinite programming
relaxation are particularly efficient, because they explore the
structure of the MAX-CUT problem.
\end{proof}
```

Other theorem-like environments (theorem, definition, remark, etc.) need to be defined as required, e.g. using \newtheorem{theorem}{Theorem} in the preamble of your .tex file (see the preamble of interactcadsample.tex for more examples). You can define the numbering scheme for these structures however suits your article best. Please note that the format of the text in these environments may be changed if necessary to match the style of individual journals by the typesetter during preparation of the proofs.

8.7. *Mathematics*

8.7.1. *Displayed mathematics*

The interact class file will set displayed mathematical formulas centred on the page without equation numbers if you use the displaymath environment or the equivalent \[...\] construction. For example, the equation

$$\hat{\theta}_{w_i} = \hat{\theta}(s(t, \mathcal{U}_{w_i}))$$

was typeset using the commands

```
\[
 \hat{\theta}_{w_i} = \hat{\theta}(s(t,\mathcal{U}_{w_i}))
\]
```

For those of your equations that you wish to be automatically numbered sequentially throughout the text for future reference, use the equation environment, e.g.

$$\hat{\theta}_{w_i} = \hat{\theta}(s(t, \mathcal{U}_{w_i})) \tag{1}$$

was typeset using the commands

```
\begin{equation}
 \hat{\theta}_{w_i} = \hat{\theta}(s(t,\mathcal{U}_{w_i}))
\end{equation}
```

Part numbers for sets of equations may be generated using the subequations environment, e.g.

$$\varepsilon \rho w_{tt}(s,t) = N[w_s(s,t), w_{st}(s,t)]_s, \tag{2a}$$

$$w_{tt}(1,t) + N[w_s(1,t), w_{st}(1,t)] = 0, \tag{2b}$$

which was typeset using the commands

```
\begin{subequations} \label{subeqnexample}
\begin{equation}
    \varepsilon \rho w_{tt}(s,t) = N[w_{s}(s,t),w_{st}(s,t)]_{s},
    \label{subeqnparta}
\end{equation}
\begin{equation}
    w_{tt}(1,t)+N[w_{s}(1,t),w_{st}(1,t)] = 0,    \label{subeqnpartb}
\end{equation}
\end{subequations}
```

Displayed mathematics should be given end-of-line punctuation appropriate to the running text sentence of which it forms a part, if required.

8.7.2.   Math fonts

8.7.2.1.   Superscripts and subscripts. Superscripts and subscripts will automatically come out in the correct size in a math environment (i.e. enclosed within \(...\) or $...$ commands in running text, or within \[...\] or the equation environment for displayed equations). Sub/superscripts that are physical variables should be italic, whereas those that are labels should be roman (e.g. C_p, T_{eff}). If the subscripts or superscripts need to be other than italic, they must be coded individually.

8.7.2.2.   Upright Greek characters and the upright partial derivative sign. Upright lowercase Greek characters can be obtained by inserting the letter u in the control code for the character, e.g. \umu and \upi produce µ (used, for example, in the symbol for the unit microns – µm) and π (the ratio of the circumference of a circle to its diameter). Similarly, the control code for the upright partial derivative ∂ is \upartial. Bold lowercase as well as uppercase Greek characters can be obtained by {\bm \gamma}, for example, which gives $\boldsymbol{\gamma}$, and {\bm \Gamma}, which gives $\boldsymbol{\Gamma}$.

Acknowledgement(s)

An unnumbered section, e.g. \section*{Acknowledgements}, may be used for thanks, etc. if required and included *in the non-anonymous version* before any Notes or References.

Disclosure statement

An unnumbered section, e.g. \section*{Disclosure statement}, may be used to declare any potential conflict of interest and included *in the non-anonymous version* before any Notes or References, after any Acknowledgements and before any Funding information.

Funding

An unnumbered section, e.g. \section*{Funding}, may be used for grant details, etc. if required and included *in the non-anonymous version* before any Notes or References.

Notes on contributor(s)

An unnumbered section, e.g. \section*{Notes on contributors}, may be included *in the non-anonymous version* if required. A photograph may be added if requested.

Nomenclature/Notation

An unnumbered section, e.g. \section*{Nomenclature} (or \section*{Notation}), may be included if required, before any Notes or References.

Notes

An unnumbered Notes section may be included before the References (if using the endnotes package, use the command \theendnotes where the notes are to appear, instead of creating a \section*).

9. References

9.1. *References cited in the text*

9.2. *The list of references*

References should be listed at the end of the main text in alphabetical order by authors' surnames, then chronologically (earliest first). If references have the same author(s), editor(s), etc., arrange by year of publication, with undated works at the end. A single-author entry precedes a multi-author entry that begins with the same name. If the reference list contains two or more items by the same author(s) in the same year, add a, b, etc. and list them alphabetically by title. Successive entries by two or more authors when only the first author is the same are alphabetized by co-authors' surnames. If a reference has more than ten named authors, list only the first seven, followed by 'et al.'. If a reference has no author or editor, order by title; if a date of publication is impossible to find, use 'n.d.' in its place.

The following list shows some sample references prepared in the Taylor & Francis Chicago author-date style.

(Adelman 2009; Albiston 2005)

10. APPENDIX: R Code for perming data analysis

```
library(dplyr)
library(stringr)
library(tidyverse)
library(tidyr)
library(dlookr)
library(table1)
library(alphaOutlier)
library(boot)
```

```
library(htmlTable)
library(ggplot2)
library(scales)
library(factoextra)
library(lubridate)
```

References

Adelman, Rachel. 2009. "'Such Stuff as Dreams Are Made On': God's Footstool in the Aramaic
    Targumim and Midrashic Tradition." Paper presented at the annual meeting for the Society
    of Biblical Literature, New Orleans, Louisiana, Nov. 21—24.
Albiston, Catherine R. 2005. "Bargaining in the Shadow of Social Institutions: Competing
    Discourses and Social Change in the Workplace Mobilization of Civil Rights." *Law and
    Society Review* 39 (1): 11–47.

11.   Appendices

Any appendices should be placed after the list of references, beginning with the command \appendix followed by the command \section for each appendix title, e.g.

```
\appendix
\section{This is the title of the first appendix}
\section{This is the title of the second appendix}
```

produces:

Appendix A. This is the title of the first appendix

Appendix B. This is the title of the second appendix

Subsections, equations, figures, tables, etc. within appendices will then be automatically numbered as appropriate. Some theorem-like environments may need to have their counters reset manually (e.g. if they are not numbered within sections in the main text). You can achieve this by using \numberwithin{remark}{section} (for example) just after the \appendix command.

Please note that if the endfloat package is used on a document containing appendices, the \processdelayedfloats command must be included immediately before the \appendix command in order to ensure that the floats in the main body of the text are numbered as such.

Appendix A. Troubleshooting

Authors may occasionally encounter problems with the preparation of a manuscript using LaTeX. The appropriate action to take will depend on the nature of the problem:

(i) If the problem is with LaTeX itself, rather than with the actual macros, please consult an appropriate LaTeX 2_ε manual for initial advice. If the solution cannot be found, or if you suspect that the problem does lie with the macros, then please contact Taylor & Francis for assistance (latex.helpdesk@tandf.co.uk).

(ii) Problems with page make-up (e.g. occasional overlong lines of text; figures or tables appearing out of order): please do not try to fix these using 'hard' page make-up commands – the typesetter will deal with such problems. (You may, if you wish, draw attention to particular problems when submitting the final version of your manuscript.)

(iii) If a required font is not available on your system, allow TeX to substitute the font and specify which font is required in a covering letter accompanying your files.

Appendix B. Obtaining the template and class file

B.1. *Via the Taylor & Francis website*

This article template and the `interact` class file may be obtained via the 'Instructions for Authors' pages of selected Taylor & Francis journals.

Please note that the class file calls up the open-source LaTeX packages booktabs.sty, epsfig.sty and rotating.sty, which will, for convenience, unpack with the downloaded template and class file. The template calls for natbib.sty and subfigure.sty, which are also supplied for convenience.

B.2. *Via e-mail*

This article template, the `interact` class file and the associated open-source LaTeX packages are also available via e-mail. Requests should be addressed to `latex.helpdesk@tandf.co.uk`, clearly stating for which journal you require the template and class file.