
Automatic text summarization, 2018

Anonymous Author(s)

Affiliation

Address

email

Abstract

Today there are many documents, articles, papers and reports available in digital form. These volumes of text are invaluable sources of information and knowledge that need to be effectively summarized to be useful. In Automatic text summarization machine learning techniques are often used to generate summaries. A prior step to the generation of summaries is usually the extraction of nuggets. This paper presents the two approaches we used for the extraction of nuggets, as well as a description of their effectiveness and shortcomings.

1 Introduction

With the dramatic growth of the internet, people are overwhelmed by the tremendous amount of online information and documents. This expansion in availability of data has demanded extensive research in the automatic generation of summaries from a collection of different type of text.

Automatic summarization is the process of shortening a text document with software, in order to create a summary with the major points of the original document.

In general, there are two different approaches for text summarization: *extraction* and *abstraction*

<https://cmt.research.microsoft.com/NIPS2018/>

Please read the instructions below carefully and follow them faithfully.

1.1 Style

~15% more words in the paper compared to earlier years.

Authors are required to use the NIPS L^AT_EX style files obtainable at the NIPS website as indicated below. Please make sure you use the current files and not previous versions. Tweaking the style files may be grounds for rejection.

1.2 Retrieval of style files

The style files for NIPS and other conference information are available on the World Wide Web at

<http://www.nips.cc/>

The file `nips_2018.pdf` contains these instructions and illustrates the various formatting requirements your NIPS paper must satisfy.

The only supported style file for NIPS 2018 is `nips_2018.sty`, rewritten for L^AT_EX 2_ε. **Previous style files for L^AT_EX 2.09, Microsoft Word, and RTF are no longer supported!**

30 The L^AT_EX style file contains three optional arguments: `final`, which creates a camera-ready copy,
31 `preprint`, which creates a preprint for submission to, e.g., arXiv, and `nonatbib`, which will not
32 load the `natbib` package for you in case of package clash.

33 **New preprint option for 2018** If you wish to post a preprint of your work online, e.g., on arXiv,
34 using the NIPS style, please use the `preprint` option. This will create a nonanonymized version of
35 your work with the text “Preprint. Work in progress.” in the footer. This version may be distributed
36 as you see fit. Please **do not** use the `final` option, which should **only** be used for papers accepted to
37 NIPS.

38 At submission time, please omit the `final` and `preprint` options. This will anonymize your
39 submission and add line numbers to aid review. Please *do not* refer to these line numbers in your
40 paper as they will be removed during generation of camera-ready copies.

41 The file `nips_2018.tex` may be used as a “shell” for writing your paper. All you have to do is
42 replace the author, title, abstract, and text of the paper with your own.

43 The formatting instructions contained in these style files are summarized in Sections 2, 3, and 4
44 below.

45 **2 General formatting instructions**

46 The text must be confined within a rectangle 5.5 inches (33 picas) wide and 9 inches (54 picas) long.
47 The left margin is 1.5 inch (9 picas). Use 10 point type with a vertical spacing (leading) of 11 points.
48 Times New Roman is the preferred typeface throughout, and will be selected for you by default.
49 Paragraphs are separated by 1/2 line space (5.5 points), with no indentation.

50 The paper title should be 17 point, initial caps/lower case, bold, centered between two horizontal
51 rules. The top rule should be 4 points thick and the bottom rule should be 1 point thick. Allow 1/4 inch
52 space above and below the title to rules. All pages should start at 1 inch (6 picas) from the top of the
53 page.

54 For the final version, authors’ names are set in boldface, and each name is centered above the
55 corresponding address. The lead author’s name is to be listed first (left-most), and the co-authors’
56 names (if different address) are set to follow. If there is only one co-author, list both author and
57 co-author side by side.

58 Please pay special attention to the instructions in Section 4 regarding figures, tables, acknowledgments,
59 and references.

60 **3 Headings: first level**

61 All headings should be lower case (except for first word and proper nouns), flush left, and bold.

62 First-level headings should be in 12-point type.

63 **3.1 Headings: second level**

64 Second-level headings should be in 10-point type.

65 **3.1.1 Headings: third level**

66 Third-level headings should be in 10-point type.

67 **Paragraphs** There is also a `\paragraph` command available, which sets the heading in bold, flush
68 left, and inline with the text, with the heading followed by 1 em of space.

69 **4 Citations, figures, tables, references**

70 These instructions apply to everyone.



Figure 1: Sample figure caption.

71 4.1 Citations within the text

72 The `natbib` package will be loaded for you by default. Citations may be author/year or numeric, as
 73 long as you maintain internal consistency. As to the format of the references themselves, any style is
 74 acceptable as long as it is used consistently.

75 The documentation for `natbib` may be found at

76 `http://mirrors.ctan.org/macros/latex/contrib/natbib/natnotes.pdf`

77 Of note is the command `\citet`, which produces citations appropriate for use in inline text. For
 78 example,

79 `\citet{hasselmo}` investigated\dots

80 produces

81 Hasselmo, et al. (1995) investigated...

82 If you wish to load the `natbib` package with options, you may add the following before loading the
 83 `nips_2018` package:

84 `\PassOptionsToPackage{options}{natbib}`

85 If `natbib` clashes with another package you load, you can add the optional argument `nonatbib`
 86 when loading the style file:

87 `\usepackage[nonatbib]{nips_2018}`

88 As submission is double blind, refer to your own published work in the third person. That is, use “In
 89 the previous work of Jones et al. [4],” not “In our previous work [4].” If you cite your other papers
 90 that are not widely available (e.g., a journal paper under review), use anonymous author names in the
 91 citation, e.g., an author of the form “A. Anonymous.”

92 4.2 Footnotes

93 Footnotes should be used sparingly. If you do require a footnote, indicate footnotes with a number¹
 94 in the text. Place the footnotes at the bottom of the page on which they appear. Precede the footnote
 95 with a horizontal rule of 2 inches (12 picas).

96 Note that footnotes are properly typeset *after* punctuation marks.²

97 4.3 Figures

98 All artwork must be neat, clean, and legible. Lines should be dark enough for purposes of reproduction.
 99 The figure number and caption always appear after the figure. Place one line space before the figure

¹Sample of the first footnote.

²As in this example.

Table 1: Sample table title

Part		
Name	Description	Size (μm)
Dendrite	Input terminal	~ 100
Axon	Output terminal	~ 10
Soma	Cell body	up to 10^6

caption and one line space after the figure. The figure caption should be lower case (except for first word and proper nouns); figures are numbered consecutively.

You may use color figures. However, it is best for the figure captions and the paper body to be legible if the paper is printed in either black/white or in color.

4.4 Tables

All tables must be centered, neat, clean and legible. The table number and title always appear before the table. See Table 1.

Place one line space before the table title, one line space after the table title, and one line space after the table. The table title must be lower case (except for first word and proper nouns); tables are numbered consecutively.

Note that publication-quality tables *do not contain vertical rules*. We strongly suggest the use of the booktabs package, which allows for typesetting high-quality, professional tables:

<https://www.ctan.org/pkg/booktabs>

This package was used to typeset Table 1.

5 Final instructions

Do not change any aspects of the formatting parameters in the style files. In particular, do not modify the width or length of the rectangle the text should fit into, and do not change font sizes (except perhaps in the **References** section; see below). Please note that pages should be numbered.

6 Preparing PDF files

Please prepare submission files with paper size “US Letter,” and not, for example, “A4.”

Fonts were the main cause of problems in the past years. Your PDF file must only contain Type 1 or Embedded TrueType fonts. Here are a few instructions to achieve this.

- You should directly generate PDF files using `pdflatex`.
- You can check which fonts a PDF files uses. In Acrobat Reader, select the menu Files>Document Properties>Fonts and select Show All Fonts. You can also use the program `pdf fonts` which comes with `xpdf` and is available out-of-the-box on most Linux machines.
- The IEEE has recommendations for generating PDF files whose fonts are also acceptable for NIPS. Please see <http://www.emfield.org/icuwb2010/downloads/IEEE-PDF-SpecV32.pdf>
- `xfig` “patterned” shapes are implemented with bitmap fonts. Use “solid” shapes instead.
- The `\bbold` package almost always uses bitmap fonts. You should use the equivalent AMS Fonts:

`\usepackage{amsfonts}`

followed by, e.g., `\mathbb{R}`, `\mathbb{N}`, or `\mathbb{C}` for \mathbb{R} , \mathbb{N} or \mathbb{C} . You can also use the following workaround for reals, natural and complex:

```

135 \newcommand{\RR}{\mathbb{R}} %real numbers
136 \newcommand{\Nat}{\mathbb{N}} %natural numbers
137 \newcommand{\CC}{\mathbb{C}} %complex numbers

```

138 Note that `amsfonts` is automatically loaded by the `amssymb` package.

139 If your file contains type 3 fonts or non embedded TrueType fonts, we will ask you to fix it.

140 6.1 Margins in L^AT_EX

141 Most of the margin problems come from figures positioned by hand using `\special` or other
142 commands. We suggest using the command `\includegraphics` from the `graphicx` package.
143 Always specify the figure width as a multiple of the line width as in the example below:

```

144 \usepackage[pdftex]{graphicx} ...
145 \includegraphics[width=0.8\linewidth]{myfile.pdf}

```

146 See Section 4.4 in the graphics bundle documentation ([http://mirrors.ctan.org/macros/](http://mirrors.ctan.org/macros/latex/required/graphics/grfguide.pdf)
147 [latex/required/graphics/grfguide.pdf](http://mirrors.ctan.org/macros/latex/required/graphics/grfguide.pdf))

148 A number of width problems arise when L^AT_EX cannot properly hyphenate a line. Please give LaTeX
149 hyphenation hints using the `\-` command when necessary.

150 Acknowledgments

151 Use unnumbered third level headings for the acknowledgments. All acknowledgments go at the end
152 of the paper. Do not include acknowledgments in the anonymized submission, only in the final paper.

153 References

154 References follow the acknowledgments. Use unnumbered first-level heading for the references. Any
155 choice of citation style is acceptable as long as you are consistent. It is permissible to reduce the font
156 size to `small` (9 point) when listing the references. **Remember that you can use more than eight**
157 **pages as long as the additional pages contain *only* cited references.**

158 [1] Alexander, J.A. & Mozer, M.C. (1995) Template-based algorithms for connectionist rule extraction. In
159 G. Tesauero, D.S. Touretzky and T.K. Leen (eds.), *Advances in Neural Information Processing Systems 7*, pp.
160 609–616. Cambridge, MA: MIT Press.

161 [2] Bower, J.M. & Beeman, D. (1995) *The Book of GENESIS: Exploring Realistic Neural Models with the*
162 *GENeral NEural Simulation System*. New York: TELOS/Springer–Verlag.

163 [3] Hasselmo, M.E., Schnell, E. & Barkai, E. (1995) Dynamics of learning and recall at excitatory recurrent
164 synapses and cholinergic modulation in rat hippocampal region CA3. *Journal of Neuroscience* **15**(7):5249-5262.