Instructions for NAACL HLT 2016 Proceedings[[1]](#footnote-2)\*

|  |
| --- |
|  |
|  |
|  |
| Anonymous NAACL submission |
|  |
|  |
|  |
|  |
|  |

Abstract

This document contains instructions for preparing NAACL HLT 2016 submissions and camera-ready manuscripts. The document itself conforms to its own specifications, and is therefore an example of what your manuscript should look like. Papers are required to conform to all the directions reported in this document. In this Word template, the required formatting is preformatted for author use and further instructions are provided for how to use Word formatting.

Introduction

Sentiment analysis is an important application of Natural Language Processing (NLP) with widespread use in industry, where distilled information on opinion can be a significant source of value. Typically, this task is framed as a text classification problem whereby a text document is categorized within predefined sentiment labels.

[SOURCE: <https://arxiv.org/pdf/1801.07883.pdf>]

Over the past decade, deep learning has become a popular machine learning approach, yielding state-of-the-art results on such text classification tasks. The input features to these deep learning models are generally word embeddings which encode word co-occurrence information; and by extension, some degree of semantic information [SOURCE].

However, this approach ignores explicit linguistic information traditionally used in Natural Language Processing, such as dependency parse trees and part-of-speech tags. While there are claims that neural networks learn internal representations that encode this linguistic information [SOURCE], it is not clear if the addition of such explicit linguistic information to the input features of a deep learning text classifier improve its performance.

To address this, we evaluate the impact of including part-of-speech tag and dependency parse tree information in deep learning model input features. Specifically, we evaluate changes in performance on polarity classification in the IMDB review dataset [SOURCE].

General Instructions

Manuscripts must be in two-column format. Exceptions to the two-column format include the title, as well as the authors’ names and complete addresses (only in the final version, not in the version submitted for review), which must be centered at the top of the first page (see the guidelines in Subsection 2.5), and any full-width figures or tables. Lines should be justified, with even spacing between margins (Ctrl+J). Single-spaced lines are permitted, but authors are encouraged to use spacing at Multiple, 1.02. Do not number the pages in the camera-ready version. Start all pages directly under the top margin. See the guidelines later regarding formatting the first page.

The maximum length of a submitted manuscript is eight (8) pages for the main conference (extended to 9 pages if accepted), printed single-sided, with unlimited references. See Section 3 for additional information on the maximum number of pages.

The review process is double-blind, so do not include any author information (names, addresses) when submitting a paper for review. However, you should maintain space for names and addresses so that they will fit in the final (accepted) version.

The Ruler

The NAACL HLT 2016 style defines a printed ruler which should be present in the version submitted for review. The ruler is provided in order that reviewers may comment on particular lines in the paper without circumlocution. If you are preparing a document without the provided style files, please arrange for an equivalent ruler to appear on the final output pages. The presence or absence of the ruler should not change the appearance of any other content on the page. The camera ready copy should not contain a ruler.

In this Word template, you can add the ruler to each individual page by copying it as a table from the upper left corner of the first page and then pasting the table at the end of any paragraph on each page. You may need to delete an extra empty line after pasting it. After you press **Ctrl-A+F9**, ruler numbering will update itself. (That is, press **Ctrl** key and letter **A** key simultaneously to select all text, then release these keys, then press **F9** key to update the document). You can delete the ruler by deleting the table “Ruler” at the upper left corner of each page. One way to delete the ruler is to right-click on the word “Ruler” and then select **Delete Rows**.

It is best to add the ruler to each page after you are done editing your document. If you start seeing two rulers on the same page, delete the extra instance of the ruler. If you do not see a ruler on a page, copy it from any other page and paste it onto the page where you need it. Press **Ctrl-A+F9** to update ruler numbering.

Reviewers: Note that the ruler measurements do not align well with lines in the paper — this turns out to be very difficult to do well when the paper contains many figures and equations, and, when done, looks ugly. Just use fractional references (e.g., the first line on this page is at mark 096.5), although in most cases one would expect that the approximate location will be adequate.

Electronically-available Resources

NAACL HLT provides this description in LATEX2e (naaclhlt2016.tex) and PDF format (naaclhlt2016.pdf), along with the LATEX2e style file used to format it (naaclhlt2016.sty) and an ACL bibliography style (naaclhlt2016.bst). These files are all available at naacl.org/naacl-hlt-2016. A Microsoft Word template file (naaclhlt2016.docx) is also available at the same URL. We strongly recommend the use of these style files, which have been appropriately tailored for the NAACL HLT 2016 proceedings.

Format of Electronic Manuscript

For the production of the electronic manuscript, you must use Adobe's Portable Document Format (PDF). This format can be generated from postscript files: on Unix systems, you can use ps2pdf for this purpose; under Microsoft Windows, you can use Adobe's Distiller, or if you have cygwin installed, you can use dvipdf or ps2pdf. Note that some word processing programs generate PDF which may not include all the necessary fonts (esp. tree diagrams, symbols). When you print or create the PDF file, there is usually an option in your printer setup to include none, all, or just non-standard fonts. Please make sure that you select the option of including ALL the fonts. *Before sending it, test your PDF by printing it from a computer different from the one where it was created.* Moreover, some word processor may generate very large postscript/PDF files, where each page is rendered as an image. Such images may reproduce poorly. In this case, try alternative ways to obtain the postscript and/or PDF. One way on some systems is to install a driver for a postscript printer, send your document to the printer specifying “Output to a file”, then convert the file to PDF.

For reasons of uniformity, Adobe's **Times Roman** font should be used. MS Word users may use the Times New Roman font, which is provided by default and only slightly different. Additionally, it is of utmost importance to specify the **US-Letter format** (8.5in × 11in) when formatting the paper.

Print-outs of the PDF file on US-Letter paper should be identical to the hardcopy version. If you cannot meet the above requirements about the production of your electronic submission, please contact the publication chairs above as soon as possible.

Layout

Format manuscripts with two columns to a page, following the manner in which these instructions are formatted. The exact dimensions for a page on US-Letter paper are:

* Left and right margins: 1 inch
* Top margin: 1 inch
* Bottom margin: 1 inch
* Column width: 3.15 inches
* Column height: 9 inches
* Gap between columns: 0.2 inches

Papers should not be submitted on any other paper size. Exceptionally, authors for whom it is *impossible* to format on US-Letter paper may format for *A4* paper. In this case, they should keep the *top* and *left* margins as given above, use the same column width, height and gap, and modify the bottom and right margins as necessary. Note that the text will no longer be centered.

Again, the camera-ready **should not have** a ruler, page numbers, nor a confidentiality header. If their paper is accepted, authors must remove these from their submitted document, or else use the provided naaclhlt2016\_camera\_ready.docx.

The First Page

Center the title, author name(s) and affiliation(s) across both columns (or, in the case of initial sub- mission, space for the names). Do not use footnotes for affiliations. Use the two-column format only when you begin the abstract.

**Title:** Place the title centered at the top of the first page, in a 15 point bold font. (For a complete guide to font sizes and styles, see Table 1.) Long titles should be typed on two lines without a blank line intervening. Approximately, put the title at 1in from the top of the page, followed by a blank line, then the author name(s), and the affiliation(s) on the following line. Do not use only initials for given names (middle initials are allowed). Do not format surnames in all capitals (e.g., “Mitchell,” not “MITCHELL”). The affiliation should contain the author's complete address, and if possible, an electronic mail address. Leave about 0.75in between the affiliation and the body of the first page.

**Abstract:** Type the abstract at the beginning of the first column. The width of the abstract text should be smaller than the width of the columns for the text in the body of the paper by about 0.25in on each side. Center the word **Abstract** in a 12 point bold font above the body of the abstract. The abstract should be a concise summary of the general thesis and conclusions of the paper. It should be no longer than 200 words. The abstract text should be in 10 point font.

**Text:** Begin typing the main body of the text immediately after the abstract, observing the two-column format as shown in the present document. Use 11 points for main body text. **Indent** about 0.16in when starting a new paragraph. This is accomplished with Right click, Styles, “First Line”.

Sections

**Headings:** Type and label section and subsection headings in the style shown on the present document. Use numbered sections (Arabic numerals) in order to facilitate cross references. Number subsections with the section number and the subsection number separated by a dot, in Arabic numerals. In bold, use 11 points for subsection headings, 12 points for section headings and the Abstract heading, 15 points for the title.

**Citations:** Citations within the text appear in parentheses as (Gusfield, 1997) or, if the author's name appears in the text itself, as Gusfield (1997). Append lowercase letters to the year in cases of ambiguities. Treat double authors as in (Aho and Ullman, 1972), but write as in (Chandra et al., 1981) when more than two authors are involved. Collapse multiple citations as in (Gusfield, 1997; Aho and Ullman, 1972).

**References:** References should appear under the heading **References** at the end of the document, but before any Appendices, unless the appendices contain references. Arrange the references alphabetically by first author, rather than by order of occurrence in the text. Provide as complete a reference as possible, using a consistent format, such as the one for *Computational Linguistics* or the one in the *Publication Manual of the American Psychological Association* (American Psychological Association, 1983). Authors’ full names rather than initials are preferred. You may use **standard** abbreviations for conferences[[2]](#footnote-3) and journals[[3]](#footnote-4).

**Appendices:** Appendices, if any, directly follow the text and the references (but see above). Letter them in sequence and provide an informative title: **Appendix A. Title of Appendix**.

**Acknowledgment** sections should be a last (unnumbered) section immediately before references.

Figures and Tables

**Creating:** To create new figures or tables, copy Figure 1, then replace the content and caption with the proper text. To make sure to capture all formatting, it may be wise to capture one line before and one line after the example as you copy it and then deleting the extraneous material. Color illustrations are discouraged, unless you have verified that they will be understandable when printed in black ink.

|  |
| --- |
|  |
| Figure 1: Figure caption. |

In MSWord, authors can place the graphic and its captions inside the rows of a 2×1 table (2 rows and 1 column) with invisible borders. Specify table positioning by right-clicking its handle in the upper left corner. Place the image in the center of the first row, and the caption in the center of the second row.

**Placing:** Place figures and tables in the paper near where they are first discussed, as close as possible to the top of their respective column. Wide illustrations may run across both columns and should be placed at the top of a page.

**Captions:** Provide a caption for every table and figure; number each one sequentially in the form: “**Figure 1:** Figure caption.”, “**Table 1:** Table caption.” Type the captions of the figures and tables below the body, using 9 point text. Table and Figure labels should be bold-faced.

**Numbering:** To update the numbering, press **Ctrl-A + F9**. This will update all the numbering applicable to tables, figures, equations, and headings.

**Cross-referencing:** To add a cross reference to a figure or table:

* Place the mouse pointer at the location where you wish to add the cross reference.
* Click on the **Insert** menu, (then click **Reference**), and then **Cross-reference** in the **Links** panel.
* In the **Cross-reference** dialog box, click the caption to which you are building the text reference.
* For a figure, under **Reference Type**, click **Figure**.
* Under Insert Reference To, click Only Label and Number, then click OK.
* Once the reference is in place, apply the ‘ACL Text’ font style (size 11, no bold face).
* This is a reference to Figure 1.

Equations

An example equation is shown below:

(1)

To add new equations, authors are encouraged to copy this existing equation line, and then replace with the new equation. The numbering and alignment of equation line elements is automatic. To update equation numbering, press **Ctrl-A + F9**. Note: this will only update the number to the right of the equation; to update numbering within the text you must create a cross reference.

**Cross-referencing:** To create a cross reference for an equation:

* Create a bookmark for it.
* Select the equation number to the right of the equation. Go to **Insert**, **Bookmark** (in the **Links** panel),andthen create a name for your equation. Press **Add** to create the bookmark.
* To refer back, place the mouse pointer at the location where you wish to add the cross reference.
* Go to **Insert, Cross-reference** (in the **Links** panel).In the dialogue box, select **Bookmark** and **Bookmark Text** from each dropdown list. Uncheck **Insert as Hyperlink**, then click **OK**.
* This will make it such that whenever a new equation is added, the references to the equation will be updated when **Ctrl-A + F9** is pressed.
* This an example cross-reference to Equation (1).

Footnotes

**Footnotes:** Put footnotes at the bottom of the page. They may be numbered or referred to by asterisks or other symbols.[[4]](#footnote-5) Footnotes should be separated from the text by a line,[[5]](#footnote-6) in 9 point font.

From Submitted to Camera-Ready

To remove submission formatting for the camera-ready document, delete the header and footer on the first page (double clicking in the area, then delete the text). This will remove them from the rest of the document. To remove the ruler, click on Styles, the ACL Submission Ruler dropdown, “Select All N Instance(s)”. Then press Delete on your keyboard.

Accessibility

In an effort to accommodate the color-blind (and those printing to paper), grayscale readability for all

accepted

papers is encouraged. Color is not forbidden, but authors should ensure that tables and figures do not rely solely on color to convey critical distinctions.

|  |  |  |  |
| --- | --- | --- | --- |
|  | CNN | LSTM | FF |
|  |  |  |  |
| baseline |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |

Table 1: Font guide.

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| |  |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | --- | |  | **LSTM** | | **CNN** | | **FF** | | |  | Acc. |  | Acc. |  | Acc. |  | | baseline |  |  |  |  |  |  | | pos |  |  |  |  |  |  | | pos-embed |  |  |  |  |  |  | | parse |  |  |  |  |  |  | | pos + parse |  |  |  |  |  |  | | pos-embed + parse |  |  |  |  |  |  | |
| Table 1: Test accuracy without word-embedding training |

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| |  |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | --- | |  | **LSTM** | | **CNN** | | **FF** | | |  | Acc. |  | Acc. |  | Acc. |  | | baseline |  |  |  |  |  |  | | pos |  |  |  |  |  |  | | pos-embed |  |  |  |  |  |  | | parse |  |  |  |  |  |  | | pos + parse |  |  |  |  |  |  | | pos-embed + parse |  |  |  |  |  |  | |
| Table 2: Test accuracy with word-embedding training |

Length of Submission

The NAACL HLT 2016 main conference accepts long and short paper submissions. Long papers may consist of up to eight (8) pages of content, plus unlimited pages for references. Accepted long papers will be given one additional page (up to 9 pages with unlimited pages for references) so that reviewers’ comments can be taken into account. Short papers may consist of up to four (4) pages of content, plus unlimited pages for references. Accepted short papers will be given five (5) pages in the proceedings and unlimited pages for references. For both long and short papers, all illustrations and appendices must be accommodated within these page limits, observing the formatting instructions given in the present document.

Double-blind review process

As the reviewing will be blind, the paper must not include the authors’ names and affiliations. Further- more, self-references that reveal the author’s identity, e.g., “We previously showed (Smith, 1991) ...” must be avoided. Instead, use citations such as “Smith previously showed (Smith, 1991) ...” Papers that do not conform to these requirements will be rejected without review. In addition, please do not post your submissions on the web until after the review process is complete (in special cases this is permitted: see the multiple submission policy below).

We will reject without review any papers that do  
not follow the official style guidelines, anonymity  
conditions and page limits.

Multiple Submission Policy

Papers that have been or will be submitted to other meetings or publications must indicate this at submission time. Authors of papers accepted for presentation at NAACL HLT 2016 must notify the program chairs by the camera-ready deadline as to whether the paper will be presented. All accepted papers must be presented at the conference to appear in the proceedings. We will not accept for publication or presentation papers that overlap significantly in content or results with papers that will be (or have been) published elsewhere.

Preprint servers such as arXiv.org and ACL-related workshops that do not have published proceedings in the ACL Anthology are not considered archival for purposes of submission. Authors must state in the online submission form the name of the workshop or preprint server and title of the non-archival version. The submitted version should be suitably anonymized and not contain references to the prior non-archival version. Reviewers will be told: “The author(s) have notified us that there exists a non-archival previous version of this paper with significantly overlapping text. We have approved submission under these circumstances, but to preserve the spirit of blind review, the current submission does not reference the non-archival version.” Reviewers are free to do what they like with this information.

|  |
| --- |
|  |

Authors submitting more than one paper to NAACL HLT must ensure that submissions do not overlap significantly (> 25%) with each other in content or results. Authors should not submit short and long versions of papers with substantial overlap in their original contributions.

STREAM Tools

This Microsoft Word file has been preset for compatible use with the STREAM Tools template designed for creating well-formatted reports and papers with Microsoft Word. The principles behind this template and others STREAM templates are explained in (Mamishev, 2010; Mamishev, 2013).

Acknowledgments

Do not number the acknowledgment section.

References

Alfred. V. Aho and Jeffrey D. Ullman. 1972. *The Theory of Parsing, Translation and Compiling, volume 1*. Prentice-Hall, Englewood Cliffs, NJ.

American Psychological Association. 1983. *Publications Manual.* American Psychological Association, Washington, DC.

Association for Computing Machinery. 1983. *Computing Reviews*, 24(11):503-512.

Ashok K. Chandra, Dexter C. Kozen, and Larry J.Stockmeyer. 1981. Alternation. *Journal of the Association for Computing Machinery*, 28(1):114-133.

Dan Gusfield. 1997. *Algorithms on Strings, Trees and Sequences*. Cambridge University Press, Cambridge, UK.

Alexander V. Mamishev and Murray Sargent. 2013. *Creating Research and Scientific Documents Using Microsoft Word*. Microsoft Press, Redmond, WA.

Alexander V. Mamishev and Sean D. Williams. 2010. *Technical Writing for Teams: The STREAM Tools Handbook*. Wiley-IEEE Press, Hoboken, NJ.

# References

Biagioni, R. (2016). *The SenticNet Sentiment Lexicon: Exploring Semantic Richness in Multi-Word Concepts.* Springer.

Honnibal, M. A. (2017). spaCy 2: Natural language understanding with Bloom embeddings, convolutional neural networks and incremental parsing.

Jacob Andreas, D. K. (2014). How much do word embeddings encode about syntax? *ACL*.

Maas, A. L. (2011). Learning Word Vectors for Sentiment Analysis. *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies* (pp. 142--150). Portland: Association for Computational Linguistics.

O Das, R. B. (2014). Sentiment Analysis of Movie Reviews using POS tags. *International Journal of Computer Applications*, 36-41.

Rui Liu, J. H. (2018). *Structural Embedding of Syntactic Trees for Machine Comprehension.* Carnegie Mellon University.

Seyed Mahdi Rezaeinia, A. G. (2018). *Text Classification based on Multiple Block.*

Terra Blevins, O. L. (2018). Deep RNNs Encode Soft Hierarchical Syntax. *ACL*.

1. \* This document has been adapted from the instructions for earlier ACL and NAACL proceedings, including those for NAACL-HLT-15 by Matt Post and Adam Lopez, NAACL-HLT-12 by Nizar Habash and William Schuler, NAACL-HLT-10 by Claudia Leacock and Richard Wicen- towski, NAACL-HLT-09 by Joakim Nivre and Noah Smith, for ACL-05 by Hwee Tou Ng and Kemal Oflazer, for ACL-02 by Eugene Charniak and Dekang Lin, and earlier ACL and EACL formats. Those versions were written by several people, including John Chen, Henry S. Thompson and Donald Walker. Additional elements were taken from the formatting instructions of the *International Joint* *Conference on Artificial Intelligence*. Microsoft Word formatting was added by Alexander Mamishev (Mamishev, 2013). [↑](#footnote-ref-2)
2. https://en.wikipedia.org/wiki/  
   List\_of\_computer\_science\_conference\_acronyms [↑](#footnote-ref-3)
3. http://www.abbreviations.com/jas.php [↑](#footnote-ref-4)
4. This is how a footnote should appear. [↑](#footnote-ref-5)
5. Note the line separating the footnotes from the text. [↑](#footnote-ref-6)