Submission and Formatting Instructions for the Twenty-Sixth   
International Conference on Machine Learning (ICML 2009)

**Keywords:** boring formatting information, online learning, information extraction, robotics, computer vision

Abstract[[1]](#footnote--1)

ICML 2009 full paper submissions are due January 26, 2009. Reviewing will be blind to the identities of the authors, and therefore identifying information should not appear in any way in papers submitted for review. Submissions must be in PDF or Postscript, 8 page length limit.

Electronic Submission

As in the past few years, ICML 2009 will rely exclusively on electronic formats for submission and review. We assume that all authors will have access to standard software for word processing, electronic mail, and web file transfer.

* 1. Templates for Papers

Electronic templates for producing papers for submission are available for LaTeX and Microsoft Word. Templates are accessible on the World Wide Web at: http://icml2009.cs.mcgill.ca/cfp.html

Send questions about these electronic templates to kiri.wagstaff@jpl.nasa.gov.

If you are using the templates, the formatting instructions below will already be enforced. The content rules you must follow are:

* The maximum paper length is 8 pages.
* Do not include author information or acknowledgments in your initial submission. Do include keywords.
* **New for 2009:** You must select an Area Chair to oversee the review process for your paper. You do not need to indicate this selection in your paper; it will be done during the online submission process.
* Place figure captions *under* the figure (and omit titles from inside the graphic file itself). Place table captions *over* the table.
* References must include page numbers whenever possible and be as complete as possible. Place multiple citations in chronological order.

Please see below for details on each of these items.

* 1. Submitting Papers

Submission to ICML 2009 will be entirely electronic, via a web site (not email). The URL and information about the submission process will appear on the conference web site at:

http://icml2009.cs.mcgill.ca/instructions.html

**Paper Deadline:** The deadline for paper submission to ICML 2009 is Monday, January 26, 2009, at 11:59 p.m. Samoa time. If your full submission does not reach us by this date, it will not be considered for publication. There is no separate abstract submission this year.

**Anonymous Submission:** To facilitate blind review, no identifying author information should appear on the title page or in the paper itself. Section 2.3 will explain the details of how to format this.

**Simultaneous Submission:** ICML will not accept any paper which, at the time of submission, is under review for another conference or a journal; is under review elsewhere; or has already been published. This policy also applies to papers that overlap substantially in technical content with papers under review or previously published. Authors are also not permitted to submit their papers elsewhere during ICML's review period.

To ensure our ability to print submissions, authors must provide their manuscripts in **postscript** or **PDF** format. Furthermore, please make sure that files contain only Type-1 fonts (e.g., using the program pdffonts in linux or using File/DocumentProperties/Fonts in Acrobat). Other fonts (like Type-3) might come from graphics files imported into the document.

If you are preparing your paper in Word, please use a generic postscript or a PDF driver to ensure its printability in other environments. Authors using **Word** must convert their document to postscript or PDF. Most of the latest versions of Word have the facility to do this automatically. Submissions will not be accepted in Word format or any format other than postscript or PDF. Really. We're not joking. Don't send Word. Those who use **LaTeX** to format their accepted papers need to pay close attention to the typefaces used. Specifically, when converting the dvi output of LaTeX to Postscript the default behavior is to use non-scalable Type-3 PostScript bitmap fonts to represent the standard LaTeX fonts. The resulting document is difficult to read in electronic form; the type appears fuzzy. To avoid this problem, dvips must be instructed to use an alternative font map. This can be achieved with something like the following commands:

**dvips -Ppdf -tletter -G0 -o paper.ps paper.dvi**

**ps2pdf paper.ps**

Note that it is a zero following the “-G”. This tells dvips to use the config.pdf file (and this file refers to a better font mapping).

Another alternative is to use the **pdflatex** program instead of straight LaTeX. This program avoids the Type-3 font problem, however you must ensure that all of the fonts are embedded (use pdffonts). If they are not, you need to configure pdflatex to use a font map file that specifies that the fonts be embedded. Also you should ensure that images are not downsampled or otherwise compressed in a lossy way.

* 1. Reacting to Reviews

We will continue the ICML tradition in which the authors are given the option of providing a short reaction to the initial reviews. These reactions will be taken into account in the discussion among the reviewers and PC-members.

* 1. Submitting Final Camera-Ready Copy

Final versions of papers accepted for publication should follow the same format and naming convention as initial submissions, except of course that the normal author information (names and affiliations) should be given. See Section 2.3.1 for details of how to format this.

The footnote, “Preliminary work. Under review by the International Conference on Machine Learning (ICML). Do not distribute.”' must be modified to “Appearing in *Proceedings of the 26th International Conference on Machine Learning*, Montreal, Canada, 2009. Copyright 2009 by the author(s)/owner(s).” For those using the LaTeX style file, simply change \usepackage{icml2009} to \usepackage[accepted]{icml2009}. Authors using Word must edit the footnote on the first page of the document themselves.

Camera-ready copies should have the title of the paper as running head on each page except the first one. The running title consists of a single line centered above a horizontal rule which is 1 point thick. The running head should be centered, bold and in 9 point type. The rule should be 10 points above the main text. For those using the LaTeX style file, the original title is automatically set as running head using the fancyhdr package which can be obtained at the ICML 2009 web site. In case that the original title exceeds the size restrictions, a shorter form can be supplied by using

\icmltitlerunning**{…}**

just before \begin{document}. Authors using Word must edit the header of the document themselves.

Approach

In tackling the problem, we focused most of our initial efforts on the feature engineering process.

* 1. Features

1. Due to the mixed nature of the problem, we explored a variety of text-based, as well as speech-based features, utilizing both traditional features of NER, as well as novel features inspired by the speech community.
   * 1. Bag of words

For our baseline model, we use the simple bag-of-word feature. This feature simply includes the surface form of all the words inside a window of fixed size around each word. This is a simple feature, which has proved to work well in Named Entity Recognition with clean text (CITATION NEEDED).

We did not perform any regularization on the text, e.g. stemming, lemmatization…, because we believe it would not help in a domain with noisy input such as ours. These techniques perhaps would help reduce the error rate of an ASR in cases where there can be confusion between possessive endings and plural forms. However, in NER on speech hypothesis, it is not clear that it would help.

* + 1. Phonetic features

In the speech domain, it is likely that an incorrect hypothesis word sounds similar to the original word. This intuition guided us to a family of features based on the phonetic structure of words. For each word in the speech hypothesis, we used an off-the-shelf text-to-phone program to find the phones that comprise it. From this sequence of phones, we extracted various features, such as phone unigrams, phone bigrams. To further de-emphasize the influence of errors caused by similar sounding phones, we devised various groupings of phones into phone classes, and then used these class names in place of the exact phone names. We used phone class sequence, “bag of class”, as well as phone class pattern as our features, where the phone class pattern feature is computed as the regular expression representing the sequence of phones.

* + 1. Base phrase chunk labels

Names always appear in noun phrases. Therefore, to detect names, we attempted to find noun phrases in the data to use as features. Quick experiments with the Stanford parser (CITATION) showed that it was quite robust to noise in the data. For example, given “in the frustrate”, it tagged as a noun the word “frustrate”, which according to WordNet (CITATION) never acts as a noun. While the Stanford parser is a regular syntactic parser and gives syntax labels for each word, we used it as a shallow parser, stopping at the base noun phrase level. The feature is a binary one, which distinguishes whether a word appears in a noun phrase or not.

* + 1. Other features

Other features suitable to use on a speech output corpus include acoustic confidence and language model confidence. Due to a problem with our data, which will be explained in section 3.1, we were not able to use these features. However, for a similar experiment with more data, it is worth trying.

*Figure 1*. Historical locations and numberof accepted papers for International Machine Learning Conferences (ICML 1993 ⎯ ICML 2008) and International Workshops on Machine Learning (ML 1988 – ML 1992). At the time this figure was produced, the number of accepted papers for ICML 2008 was unknown and instead estimated.

We should also mention some common features that we chose not to use. Part-of-speech tags, character n-grams and word shape are such features. With our noisy data, we decided to exclude these features to mitigate potential damage caused by not having the correct text from which to extract them.

* 1. Learning models

1. We used Conditional Random Fields (Lafferty, McCallumn, Perreira, 2001), which is the most popular model in sequential labeling problems, as the learning model with which to experiment the features. Afterwards, we picked the combination of features that worked best in CRFs and used them with a variety of other sequential learners for a better perspective. We experimented with Conditional Markov Models using probabilistic SVM (SVMCMM) (CITATION), MEMM (McCallum, Freitag, 2000), as well as Voted Perceptron HMM (CITATION).

Evaluation

* 1. Partitioning the Text

You should organize your paper into sections and paragraphs to help readers place a structure on the material and understand its contributions.

* + 1. Sections and Subsections

Section headings should be numbered, flush left, and set in 11 pt bold type with the content words capitalized. Leave 0.25 inches of space before the heading and 0.15 inches after the heading.

Similarly, subsection headings should be numbered, flush left, and set in 10 pt bold type with the content words capitalized. Leave 0.2 inches of space before the heading and 0.13 inches afterward.

Finally, subsubsection headings should be numbered, flush left, and set in 10 pt small caps with the content words capitalized. Leave 0.18 inches of space before the heading and 0.1 inches after the heading. Please use no more than three levels of headings.

* + 1. Paragraphs and Footnotes

Within each section or subsection, you should further partition the paper into paragraphs. Do not indent the first line of a given paragraph, but insert a blank line between succeeding ones.

You can use footnotes [[2]](#footnote-0) to provide readers with additional information about a topic without interrupting the flow of the paper. Indicate footnotes with a number in the text where the point is most relevant. Place the footnote in 9 point type at the bottom of the column in which it appears. Precede the first footnote in a column with a horizontal rule of 0.8 inches.[[3]](#footnote-1)

* 1. Figures

You may want to include figures in the paper to help readers visualize your approach and your results. Such artwork should be centered, legible, and separated from the text. Lines should be dark and at least 0.5 points thick for purposes of reproduction, and text should not appear on a gray background.

Label all distinct components of each figure. If the figure takes the form of a graph, then give a name for each axis and include a legend that briefly describes each curve. Do not include a title inside the figure; instead, the caption should serve this function.

Number figures sequentially, placing the figure number and caption *after* the graphics, with at least 0.1 inches of space before the caption and 0.1 inches after it, as in Figure 1. The figure caption should be set in 9 point type and centered unless it runs two or more lines, in which case it should be flush left. You may float figures to the top or bottom of a column, and you may set wide figures across both columns (use the environment figure\* in LaTeX), but always place two-column figures at the top or bottom of the page.

* 1. Algorithms

If you are using LaTeX, please use the “algorithm” and “algorithmic” environments to format pseudocode. These require the corresponding stylefiles, algorithm.sty and algorithmic.sty, which are supplied with this package. Algorithm 1 shows an example.

**Algorithm 1** Bubble Sort

**Input**: data *x* , size *m*

**repeat**

Initialize *noChange = true*.

**for**  *i = 1* to m-1

**if** *xi*  > xi+1

Swap *xi*  and xi+1

*noChange = false*

**end if**

**end for**

**until**  *noChange* is *true*

* 1. Tables

You may also want to include tables that summarize material. Like figures, these should be centered, legible, and numbered consecutively. However, place the title *above* the table with at least 0.1 inches of space before the title and the same after it, as in Table1. The table title should be set in 9 point type and centered unless it runs two or more lines, in which case it should be flush left.

Tables contain textual material that can be typeset, as contrasted with figures, which contain graphical material that must be drawn. Specify the contents of each row and column in the table's topmost row. Again, you may float tables to a column's top or bottom, and set wide tables across both columns, but place two-column tables at the top or bottom of the page.

*Table 1*. Classification accuracies for naive Bayes and flexible Bayes on various data sets.

|  |  |  |  |
| --- | --- | --- | --- |
| Data Set | Naïve | Flexible | Better? |
| Breast | 95.9 ± 0.2 | 96.7 ± 0.2 | ✓ |
| Cleveland | 83.3 ± 0.6 | 80.0 ± 0.6 | x |
| Credit | 74.8 ± 0.5 | 78.3 ± 0.6 |  |
| Glass2 | 61.9 ± 1.4 | 83.8 ± 0.7 | ✓ |
| Horse | 73.3 ± 0.9 | 69.7 ± 1.0 | x |
| Meta | 67.1 ± 0.6 | 76.5 ± 0.5 | ✓ |
| Pima | 75.1 ± 0.6 | 73.9 ± 0.5 |  |
| Vehicle | 44.9 ± 0.6 | 61.5 ± 0.4 | ✓ |

* 1. Citations and References

Authors should cite their own work in the third person in the initial version of their paper submitted for blind review.

Please use APA reference format regardless of your formatter or word processor. If you rely on the LaTeX bibliographic facility, use mlapa.sty and mlapa.bst at the ICML 2009 web site to obtain this format.

Citations within the text should include the authors' last names and year. If the authors' names are included in the sentence, place only the year in parentheses, for example when referencing Rob Schapire's seminal result (1990). Otherwise place the entire reference in parentheses with the authors and year separated by a comma (Schapire, 1990). You can anonymize the bibliographic entries during submission, as in (Authors, 1900), if you believe the full citation would compromize the anonymous nature of the submission.

List multiple references separated by semicolons (Kearns, 1989; Schapire, 1990; Neal, 1993). Use the “et al.” construct only for citations with four or more authors or after listing all authors to a publication in an earlier reference.

Use an unnumbered first-level section heading for the references, and use a hanging indent style, with the first line of the reference flush against the left margin and subsequent lines indented by 10 points. The references at the end of this document give examples for journal articles, conference publications, book chapters, books, edited volumes, technical reports, and dissertations.

Alphabetize references by the surnames of the first authors, with single author entries preceding multiple author entries. Order references for the same authors by year of publication, with the earliest first.

Acknowledgments

**Do not** include acknowledgements in the initial version of the paper submitted for blind review.

If a paper is accepted, the final camera-ready version can (and probably should) include acknowledgements. In this case, please place such acknowledgements in an unnumbered section at the end of the paper. Typically, this will include thanks to reviewers who gave useful comments, to colleagues who contributed to the ideas, and to funding agencies and corporate sponsors that provided financial support.

References

Authors, A. (1990). Suppressed for anonymity.

Kearns, M. J. (1989). *Computational complexity of machine learning*. Doctoral Dissertation, Department of Computer Science, Harvard University.

Neal, R. (1993). *Probabilistic inference using Markov Chain Monte Carlo methods* (Technical Report CRG-TR-93-1). Dept. of Computer Science, University of Toronto.

Schapire, R. E. (1990). The strength of weak learnability. *Machine Learning*, *5*, 197-227.

Zinkevich, M. (2003). Online convex programming and generalized infinitesimal gradient descent. *Proceedings of the International Conference on Machine Learning*.

1. Preliminary work. Under review by the International Conference on Machine Learning (ICML). Do not distribute. [↑](#footnote-ref--1)
2. For the sake of readability, footnotes should be complete sentences. However, keep them to a reasonable length (say one to three sentences). [↑](#footnote-ref-0)
3. Multiple footnotes can appear in each column, in the same order as they appear in the text, but spread them across columns and pages if possible. [↑](#footnote-ref-1)