
Project Report Template for CIS 419/519

Introduction to Machine Learning

Abstract

Use this template to write your final report for CIS 419/519. It is based on the paper template used in the International Conference on Machine Learning, one of the main machine learning conferences.

1. Final Submission

Your final submission will consist of two deliverables: (1) a final report, and (2) a set of summary slides. Remember that late days cannot be used for the final project submission.

1.1. Final Report

Your final project report can be at most 5 pages long (include all text, appendices, figures, references, and anything else), and must be written in the provided L^AT_EX template.

At a minimum your final report must describe the problem/application and motivation, survey related work, discuss your approach, and describe your results/conclusions/impact of your project. It should include enough detail such that someone else can reproduce your approach and results. For inspiration on what should be included, see the project reports available on the links provided in Section ???. You will likely end up with a better report if you start by writing a 6-7 page report and then edit it down to 5 pages of well-written and concise prose.

In addition, your report must also include a figure that graphically depicts a major component of your project (e.g., your approach and how it relates to the application, etc.). Such a summary figure makes your paper much more accessible by providing a visual counterpart to the text. Developing such a concise and clear figure can actually be quite time-consuming; I often go through around ten versions before I end up with a good final version.

After the class, we are also considering posting the final reports online so that you can read about each others

University of Pennsylvania, CIS 419/519 Course Project.
Copyright 2015 by the author(s).

work. If are okay with having your final report posted online, be sure to give us explicit permission to post it in the README file, as described in the project description.

1.2. Summary Slides

In addition to the final report, you are also required to prepare a two-slide overview of your project. Details on the summary slides are available in the project description.

2. Optional Suggestions for Your Paper and Formatting Guidance

2.1. 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, be sure to include a caption describing your figure.

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 L^AT_EX), but always place two-column figures at the top or bottom of the page.

2.2. Algorithms

If you are using L^AT_EX, 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.

2.3. 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, as in Table 1.

Algorithm 1 Bubble Sort

Input: data x_i , size m
repeat
 Initialize $noChange = true$.
 for $i = 1$ **to** $m - 1$ **do**
 if $x_i > x_{i+1}$ **then**
 Swap x_i and x_{i+1}
 $noChange = false$
 end if
 end for
until $noChange$ is $true$

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

DATA SET	NAIVE	FLEXIBLE	BETTER?
BREAST	95.9±0.2	96.7±0.2	✓
CLEVELAND	83.3±0.6	80.0±0.6	×
GLASS2	61.9±1.4	83.8±0.7	✓
CREDIT	74.8±0.5	78.3±0.6	
HORSE	73.3±0.9	69.7±1.0	×
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	✓

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.

2.4. Citations and References

Please use APA reference format regardless of your formatter or word processor. If you rely on the L^AT_EX bibliographic facility, use `natbib.sty` and `icml2014.bst` included in the style-file package 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 Arthur Samuel’s pioneering work (1959). Otherwise place the entire reference in parentheses with the authors and year separated by a comma (Samuel, 1959). List multiple references separated by semicolons (Kearns, 1989; Samuel, 1959; Mitchell, 1980). Use the ‘et al.’ construct only for citations with three or more authors or after listing all authors to a publication in an earlier reference (Michalski et al., 1983).

The references at the end of this document give examples

for journal articles (Samuel, 1959), conference publications (Langley, 2000), book chapters (Newell & Rosenbloom, 1981), books (Duda et al., 2000), edited volumes (Michalski et al., 1983), technical reports (Mitchell, 1980), and dissertations (Kearns, 1989).

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. Make sure that each reference includes all relevant information (e.g., page numbers).

Acknowledgments

If you did this work in collaboration with someone else, or if someone else (such as another professor) had advised you on this work, your report must fully acknowledge their contributions. If you received external help or assistance on this project, you must cite these sources here in the acknowledgements section. If you do not have anything to list in this section, write simply “None.”

References

Duda, R. O., Hart, P. E., and Stork, D. G. *Pattern Classification*. John Wiley and Sons, 2nd edition, 2000.

Kearns, M. J. *Computational Complexity of Machine Learning*. PhD thesis, Department of Computer Science, Harvard University, 1989.

Langley, P. Crafting papers on machine learning. In Langley, Pat (ed.), *Proceedings of the 17th International Conference on Machine Learning (ICML 2000)*, pp. 1207–1216, Stanford, CA, 2000. Morgan Kaufmann.

Michalski, R. S., Carbonell, J. G., and Mitchell, T. M. (eds.). *Machine Learning: An Artificial Intelligence Approach, Vol. I*. Tioga, Palo Alto, CA, 1983.

Mitchell, T. M. The need for biases in learning generalizations. Technical report, Computer Science Department, Rutgers University, New Brunswick, MA, 1980.

Newell, A. and Rosenbloom, P. S. Mechanisms of skill acquisition and the law of practice. In Anderson, J. R. (ed.), *Cognitive Skills and Their Acquisition*, chapter 1, pp. 1–51. Lawrence Erlbaum Associates, Inc., Hillsdale, NJ, 1981.

Samuel, A. L. Some studies in machine learning using the game of checkers. *IBM Journal of Research and Development*, 3(3):211–229, 1959.