Stochastics and Machine Learning: Project 2 — Report Template

Emanuel Schaerer Maximilan von Loesch Kolja Diehl

This is a template for your report on project 2 in stochastics and machine learning. You are free to change the sections if you want. However, stick to 2 pages in total (plus one additional page for references, if needed).
Refer to the grading scheme, and make sure that all relevant of your work are clearly described. We can only give you points for things you have described!

The quality of the report is evaluated according to the following criteria (5 points each):

- The proposed approach is motivated and clearly described.
- All relevant aspects of the experiments are clearly described.
- The results are interpreted, and the findings are consistent
- The text and graphics are of good quality.

1. Approach

Describe your idea. Mention and cite sources that you have used.

1.1. A non-standard combination...

...Of at least 2 approaches, and/or significant variation of a standard approach: 20 points for creativity

2. Experiments

The evaluation of the proposed approach will be granted a maximum of 30 points. The following criteria will be checked, and 5 points will be granted for every fulfilled criterion.*Describe how you have evaluated your method, and what conclusions you draw from the findings.*

The evaluation of the proposed approach will be granted a maximum of 30 points. The following criteria will be checked, and 5 points will be granted for every fulfilled criterion.

- *At least 2 different approaches are described (one can be a standard approach as*
- *discussed in class, e.g. the unet architecture).*
- *Relevant metrics are used for the comparison.*
- *At least 3 relevant variations / hyperparameters are evaluated.*
- *The model comparison / hyperparameter selection is done on validation set*
- *The final evaluation is done on separate test set (public leaderboard data)*
- *Cases of poor / good model performance (i.e., when does the model work well, or not)*
- *are described.*

2.1. Experimental Setup

How did you evaluate you model? What data did you use, what metrics, how long did you train, etc.

2.2. Results

Here come the results. Remember to compare your approach to a at least one other approach, and to evaluate several configurations (e.g., number of layers, or other hyperparameters) of your approach.

2.3. Conclusions

How do you interpret the results? What do you learn from them?

3. Format of the Paper

This section is not intended to appear in your report. It illustrates the usage of some commands you may need for your report.

3.1. 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.

3.1.1. 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.

3.1.2. 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[1] *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.*[2]

3.2. Figures

You may want to include figures in the paper to illustrate your approach and 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). Always place two-column figures at the top or bottom of the page.*

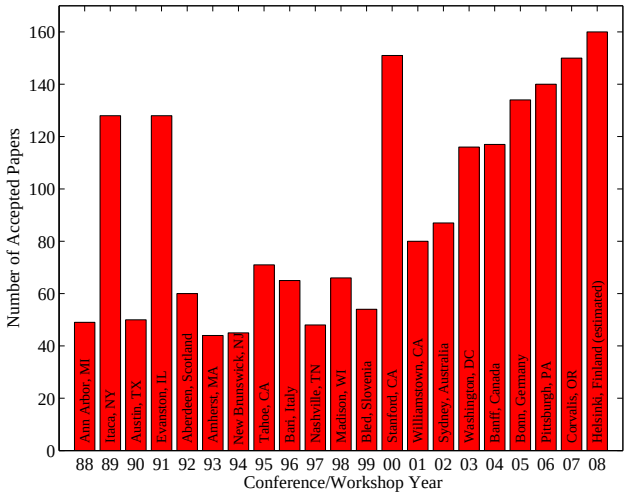

Figure 1. Historical locations and number of accepted papers for International Machine Learning Conferences (ICML 1993 – ICML 2008) and International Workshops on Machine Learning (ML 1988 – ML 1992). At the time this figure was produced, the number of accepted papers for ICML 2008 was unknown and instead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

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 $true$

3.3. 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.

3.4. 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*

[1] Footnotes should be complete sentences.

[2] 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.

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	√

at least 0.1 inches of space before the title and the same after it, as in Table 1. 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, whereas figures contain graphical material. 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. Place two-column tables at the top or bottom of the page.

3.5. Theorems and such

The preferred way is to number definitions, propositions, lemmas, etc. consecutively, within sections, as shown below.

Definition 3.1. A function $f : X \to Y$ is injective if for any $x, y \in X$ different, $f(x) \neq f(y)$.

Using Definition 3.1 we immediate get the following result:

Proposition 3.2. *If f is injective mapping a set X to another set Y, the cardinality of Y is at least as large as that of X*

Proof. Left as an exercise to the reader. □

Lemma 3.3 stated next will prove to be useful.

Lemma 3.3. *For any $f : X \to Y$ and $g : Y \to Z$ injective functions, $f \circ g$ is injective.*

Theorem 3.4. *If $f : X \to Y$ is bijective, the cardinality of X and Y are the same.*

An easy corollary of Theorem 3.4 is the following:

Corollary 3.5. *If $f : X \to Y$ is bijective, the cardinality of X is at least as large as that of Y.*

Assumption 3.6. The set X is finite.

Remark 3.7. According to some, it is only the finite case (cf. Assumption 3.6) that is interesting.

3.6. Citations and References

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).

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 (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).

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, P. (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.*

A. You *can* have an appendix here.

The `\onecolumn` *command above can be kept in place if you prefer a one-column appendix, or can be removed if you prefer a two-column appendix. Apart from this possible change, the style (font size, spacing, margins, page numbering, etc.) should be kept the same as the main body.*