
Title: Interim Report

G000 (sXXXXXXXX, sYYYYYYY, sZZZZZZZ)

Abstract

The abstract should be a few sentences (100–200 words) long, providing a concise summary of the contents of your report including the key research question(s) addressed, the methods explored, the data used, and the findings of the experiments.

1. Introduction

This document provides a template for the MLP coursework 3 interim report. This template structures the report into sections, which you may use, or you can structure it differently if you wish. If you want to use subsections within a section that is fine. In this template the text in each section will include a very brief outline of what you should include in each section, along with some practical LaTeX examples (for example figures, tables, algorithms). Your document should be no longer than **five pages**, with an additional page (or more!) allowed for references.

You should give a broad introduction to the project, including citations to related work. Your aim here is to explain why the project is addressing an interesting topic, and how it relates to things that have been done in the area.

You should make clear what are the aims and objectives of the project, what are the research questions being addressed. Be precise. In this section you should make clear what the project's contribution is: how is it different to what is already done.

The interim report should state the objectives of the project, which are related to the research questions. What experiments do you plan to carry out? You can differentiate between core objectives, and optional objectives you hope to achieve if things go well. The conclusions in your final report should relate to these objectives.

Use bibtex to organise your references – in this case the references are in the file `example-refs.bib`. Here is an example reference (Langley, 2000).

2. Data set and task

Clearly describe the data set and task you will be exploring. If the data requires any preprocessing, then explain this. The description should be in enough detail such that your work would be reproducible by another group. Describe how you will evaluate the task (for example, classification accuracy). Use citations where appropriate.

3. Methodology

Explain clearly the technical methodology, the models and algorithms that are used. Approaches that were covered in the lectures can be described briefly, but if you are using modifications to such approaches make sure these are clearly described. Again use citations to the literature. Avoid pasting or paraphrasing the methodology directly from the cited work. Instead consider why you are using these approaches, and how they can best be presented in the context of your work. Keep examples consistent with your domain, and notation consistent with your paper. Consider also how you might have adapted the approaches to work in your case.

If you present algorithms, you can use the `algorithm` and `algorithmic` environments to format pseudocode (for instance, Algorithm 1). These require the corresponding style files, `algorithm.sty` and `algorithmic.sty` which are supplied with this package.

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  $true$ 
```

4. Experiments

The interim report should include some experimental results. In most cases these will be baseline experiments. Baseline experiments refer to experiments conducted using well-understood approaches against which you can compare later results. In particular, you need to be able to justify that the baseline performance is a reasonable hurdle for your solutions to overcome. This is in contrast to a benchmark, which would be the performance you would need to beat to indicate positive findings. Baselines are used, among other things, to test that both your benchmark(s) and proposed solutions have been properly set up and trained. For example if you were exploring a new data set, the baselines might include linear networks and deep neural networks with different numbers of hidden layers; if you were exploring a different approach to regularisa-

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	✓

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

tion, then the baselines would include no regularisation, and conventional techniques such as L1, L2, and dropout. You can include the results of any further experiments in your interim report.

Present the experimental results clearly and concisely. Usually a result is in comparison or contrast to a result from another approach please make sure that these comparisons/contrasts are clearly presented. You can facilitate comparisons either using graphs with multiple curves or (if appropriate, e.g. for accuracies) a results table. You need to avoid having too many figures, poorly labelled graphs, and graphs which should be comparable but which use different axis scales. A good presentation will enable the reader to compare trends in the same graph – each graph should summarise the results relating to a particular research (sub)question.

There is no need to include code or specific details about the compute environment.

As before, your experimental sections should include graphs (for instance, figure 1) and/or tables (for instance, table 1)¹, using the figure and table environments, in which you use `\includegraphics` to include an image (pdf, png, or jpg formats). Please export graphs as `vector graphics` rather than `raster files` as this will make sure all detail in the plot is visible. Matplotlib supports saving high quality figures in a wide range of common image formats using the `savefig` function. **You should use `savefig` rather than copying the screen-resolution raster images outputted in the notebook.** An example of using `savefig` to save a figure as a PDF file (which can be included as graphics in a `LATEX` document is given in the coursework document.

If you need a figure or table to stretch across two columns use the `figure*` or `table*` environment instead of the `figure` or `table` environment. Use the `subfigure` environment if you want to include multiple graphics in a single figure.

¹These examples were taken from the ICML template paper.

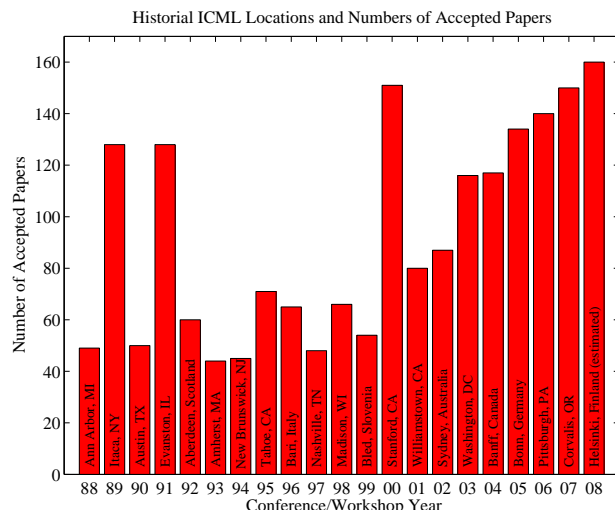


Figure 1. Historical locations and number of accepted papers for International Machine Learning Conferences (ICML 1993 – ICML 2008) and International Workshops on Machine Learning (ML 1988 – ML 1992). At the time this figure was produced, the number of accepted papers for ICML 2008 was unknown and instead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

5. Interim conclusions

What have you learned so far? Do the experiments indicate that the project is feasible? Do the experiments indicate that you should consider changes to your original plan? Can you compare your results so far to what has been reported in the literature?

6. Plan

Based on what you have done so far, present a plan for the rest of the project. Are there any changes to the objectives? What are the risks? Do you need a backup plan?

References

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. URL <http://papers.nips.cc/paper/4824-imagenet-classification-with-deep-convolutional-neural-networks.pdf>.