How to write a report

Simone Teufel

Michaelmas 2018/19
Report: Goal

- To describe your work as a reimplementation of Pang et al. (2002)
- To explain in detail how you followed the instructions
- To mimic the language and organisation of a research paper
Report: Structure

- Introduction
- Background
- Method (reimplementation choices)
- Result (with some discussion)
- Conclusion
General Tips

- Typeset in two column
- Use latex if you can
- Math notation – define each variable (either in running text, or in a pseudo-legenda after or before the equation)
- Avoid colloquial language – everything can be said in a scientific-sounding way
- Avoid lengthy sequences of actions you did in favour of results / functionality of algorithm. If sequence is necessary give main idea first
- In each paragraph: say the main idea first
Marking-specific tips

- Allocation of space in paper should mirror your (perceived) effort
  - Do not spend space on “obvious” things
  - Spend more space on things that make your solution stand out
  - Or where you spent more effort than expected / than others
- If you don’t write it, we cannot give you marks for it
Introduction

- Here: quite short
- Phrase as a replication experiment
- State that you were “given the data in the framework of a course in NLP”
- Describe alternative systems; e.g., NB unsmoothed – NB smoothed – SVM
- Define technical terminology you will need (later)
Special case here: reimplementation
Entire Background section reserved for Pang et al.
Introduce all ideas they had first here (because of the timeline of discovery).
Do not (in some later section) present anything they already did, as if you invented it; you didn’t.
The only interesting things to say about your reimplementation is if there was a change or discovery. Otherwise just say “we reimplemented this system using XXX programming language...”
More generally, this is the “My Method” section

Be specific – the reader needs details (in general: to reimplement your work; here – in order to analyse your numbers and mark your effort)

Give “intermediate stage results”, e.g., After eliminating all features which occurred less than 2 times, 3,289 features remained
Results

- Tabulate your results in stages; normally in more than one table
- Each table should have a theme (e.g., comparison between symbolic methods, comparison between different ML methods in 2 tables)
- Table caption should be descriptive of the results and the exact version of dataset you are using
- Metric should be clear from table even without reader having to read the text to find out
- Text should contain main information without having to look at tables. Tables should also be stand-alone.
- Text contains pointers to tables (and probably repeat main numbers).
- Report numerical results with what can be reasonably thought to be significant digits.
- Indicate significance (triangular matrix or summary of trends, if possible and/or useful for your message) in tables.
- Often enough to say word “significant” only once in text.
- First time to state the word “significant”, describe/state test in footnote (e.g., significance level, name/type of test, two- or one-tailed).
Interpretation of results

- Main result first
- The one that corresponds to your main hypothesis
- It worked or it didn’t work
- Then maybe: impact of features; ablation tests or feature selection results
- Your comparison ground – baselines and competitor systems; maybe ceilings.
- Notion of “interestingness” of a result – can you connect the result to a related observation that might be slightly non-obvious
- Later in your research careers: Cross-links to other people’s results
Null results

- If you find a positive effect of your intervention, it obviously worked.
- Something else may have hypothetically worked better, but it does not matter because you brought positive proof for your intervention.
- If you find a negative effect of your intervention, it may have two reasons: you didn’t try hard enough, or the effect is really not there.
- If you believe the latter, you have to convince your readers that you tried everything reasonable.
- Not a problem in educational context, but often so in the “real science world”.

Simone Teufel
How to write a report
An ideal report

- Precise, scientific-sounding, technical, to the point
- Little general “waffle”/chit-chat
- Not boring – because you don’t explain obvious things too much
- Efficient delivery of (only) the facts that we need to know to understand/reimplement
- Results visually well-presented and described with the correct priority of importance of sub-results
- Insightful analysis – speculation should connect to something interesting and not be too much; the reader “learns something new”
- No typos, no colloquialisms – well-considered language
- This normally means several re-draftings (re-orderings of information)
Thank you!