L90 Practical, Report on Baseline System

Overview of Natural Language Processing

Simone Teufel

2016/11/11

Report: Baseline System for Sentiment Classification

- Write a two-page report describing your baseline system
- Submit to Student Admin next Friday
- I will give you feedback by email by Nov 25 (our next meeting)
- Having a reasonable baseline system will allow you to judge your intervention properly
- Having practiced report writing will improve your final report

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

Structure

- Intro
- Background
- Method (reimplementation choices)
- Result (with some discussion if possible)
- Conclusion – maybe not necessary here for length and content reasons

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)
- Use "I" – you can do this up to an including your PhD thesis
- 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, I cannot give you marks for it

Intro

- Here: quite short
- Phrase as a replication experiment
- State that you were "given the data in the framework of a MPhil course in NLP"
- Describe alternatives; e.g., Symbolic – ML (as if it were initially unknown)
- Define Technical terminology you need here (maybe this task is too easy to need much)

Background

- 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 don't exist (yet).

My implementation

- More generally, this is the "My Method" section
- Be specific – in order to analyse your numbers, I need to know the details. In general – in order to reimplement your work, we need details
- Give "intermediate stage results", eg
  After eliminating all features which occurred less than 2 times, 3289 features remained.
- For instance, give info about your tokenisation method (how implemented, how many rules, how many special cases (if this is what you do)).

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)
- Metric should be clear from table even without having to read the text

Significance

- Report numerical results with what can be reasonably thought to be **significant digits**
- Indicate significance (triangular matrix or shortcuts, 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.

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
- 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 worked better, but it does not matter because you brought positive proof.
- 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.
- You believe the latter, but you have to convince your readers that you tried everything reasonable
- That is why in the "real science world", null results can be problematic.
- Here on the Mphil, they are not.

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. (which would make you sound naive)
- Efficient delivery of (only) the facts that I need to know to understand/reimplement
- Results visually well-presented and described with the correct priority of importance of sub-results
- Analysis "insightful" – 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)