CS380 Information Retrieval

HW 3

Written and Oral report guidelines

WRITTEN REPORT
Due March 16, in class
1 report per team (be sure to list all team members in report)
Length: 3-10 pages. This will vary depending on presence of tables, etc.
I do not expect that the textual will exceed 2000 words (but if you find
that the report is longer, that is OK). That said, avoid writing “War and
Peace”

This report should follow the general format of a scholarly research
paper. As such it might be expected to have some or all of the following
sections:
1. Introduction/Background: what is the problem, why is the problem a
   problem?
   1-2 paragraphs
2. Previous Work: What have other people done on this problem. For
   some projects, the textbook will be a significant help. (Also the
   secondary text has some sections that apply directly.) If you have
   chosen to implement and test a technique proposed in another paper,
   some discussion of why that paper (google scholar will help with these
   sort of questions). If you are working an algorithm that proposes to fix
   an issue in another algorithm, you should discuss the problematic
   algorithm. (Although you could choose to put such discussion into the
   Introduction.) You need not cover any previous work already discussed
   in class.
   1-2 paragraphs
3. Algorithm: The algorithm you implemented at a very high level. If
   there is pseudocode, it should be very pseudo. Here describe all
   design decisions you made, the effect of those decisions, and a prima
   facie case for why those decisions are reasonable.
   For instance, “To reduce the complexity of the problem, I randomly
   excluded 75% of the vocabulary. This increased the speed of the
   retrieval by a factor of 16. The decision is reasonable because most
   words do not matter that much.” While this example covers each of
   the points I suggested, it is very brief and the prima facie case is
weak, at best.
To the extent that you made decisions, you will probably need some
supporting experimental evidence of their validity in the remainder of
the paper.
If your method used software that you did not write, it should
probably be mentioned here. (And certainly should be mentioned
somewhere.)
As long as it needs to be, this could be as much as 2/3 of the paper.
4. Methodology: How you when about testing you system. For many this
will be fairly short as you testing will be along the lines of “here are a
set of queries ...”. If that is the case, in this section you should discuss
how you came up with the set of testing queries. For others this
section might be omitted as the statement “I looked for interesting
before / after differences” is quite boring.
0-2 paragraphs
5. Results/Analysis: What you get when you apply your testing
methodology to your algorithm. For some this will look similar to the
annotated script from HW2. Realistically, this is the meat of the paper.
Of indeterminate length. If you are annotating scripts, then this could
be quite long.
6. Conclusions: Did the system do what you expected? Did it work? What
you did going forward.

ORAL REPORT

10-15 minutes
1-2 slides per section of the written report. For most presentations you
can omit the related work. (At most you might mention related work in
the algorithm (or results) discussion. For instance, “XXX did this
instead.”) The report should focus on the algorithm you implemented and
how it works. Keep in mind that for oral presentation your discussion of
the algorithm must be very simple, tending towards diagrams with circles
and arrows. Results should be very carefully chosen to illustrate specific
pointss. Inevitably you will need to NOT discuss a lot of interesting things
in your report. Do not try to get everything in by just talking very fast.