**Name: Your Name**

**Andrew ID: Your Andrew Id**

**Machine Learning for Text Mining**

**Homework 1 - Template**

1. **Statement of Assurance**
2. **Experiments**
3. Describe the custom weighing scheme that you have implemented. Explain your motivation for creating this weighting scheme.
4. Report of the performance of the 9 approaches.

I. Metric: MAP

|  |  |  |  |
| --- | --- | --- | --- |
| Method \ Weighting Scheme | NS | WS | CM |
| GPR |  |  |  |
| QTSPR |  |  |  |
| PTSPR |  |  |  |

II. Metric: Precision at 11 standard recall levels  
(Use one table for each recall level, so totally there would be 11 tables.)

|  |  |  |  |
| --- | --- | --- | --- |
| Method \ Weighting Scheme | NS | WS | CM |
| GPR |  |  |  |
| QTSPR |  |  |  |
| PTSPR |  |  |  |

III. Metric: Wall-clock running time in seconds

|  |  |  |  |
| --- | --- | --- | --- |
| Method \ Weighting Scheme | NS | WS | CM |
| GPR |  |  |  |
| QTSPR |  |  |  |
| PTSPR |  |  |  |

IV. Parameters

1. Compare these 9 approaches based on the various metrics described above.
2. Analyze these various algorithms, parameters, and discuss your general observations about using PageRank algorithms.
3. 1. What could be some novel ways for search engines to estimate whether a query can benefit from personalization?

2. What could be some novel ways of identifying the user’s interests (e.g. the user’s topical interest distribution Pr(t|u)) in general?

1. **Details of the software implementation**
2. Describe your design decisions and high-level software architecture;
3. Describe major data structures and any other data structures you used for speeding up the computation of PageRank;
4. Describe any programming tools or libraries and programming environment used;
5. Describe strengths and weaknesses of your design, and any problems that your system encountered