

CMSC 723

Project Proposal

Due: November 11, 2013

1 Proposal Assignment

Construct a proposal according to the layout here. Turn in your proposal at this link:
http://inclass.umiacs.umd.edu/perl/handin.pl?course=cmsc723_2013

Sign up for a time to meet with Andy here: <http://doodle.com/z36g7z29ugz6igx3>

Late days cannot be used for any project-related deliverable.

2 Feature Set

In this section, lay out a set of features you plan to extract from the text and available metadata. For now, just assume you have the myPersonality data – that is, everything is legal to use except for the Big Five personality data itself. For instance:

- Unigram Counts
- Timestamps
- ...

You should also include a one to two sentence blurb for each feature explaining why you think it may be useful in predicting personality in this context.

3 Plans for Extraction

In this section, lay out how you plan to extract each feature from the data. This does not have to be exhaustive and too in-depth, but you should be able to communicate why each feature is feasible for you to extract for this term project. This is the proper place to touch on established toolkits you may use (i.e. Mallet for LDA, etc.). Your goal should be to explain any particularly intriguing features – if you are somehow pulling out an author's sex, don't leave us scratching our heads as to how.

4 Development Plan

Here, describe how you plan to develop and improve your model. There is currently no test data available to you – how do you plan to work around this and still verify that the system you are building is working? Methodology for splitting data into training/development set splits belongs here, as well as any additional tools you plan to use to evaluate your system.

4.1 Deliverable Timeline

In order to help stay on track, part of your development plan should include a notion of a timeline for deliverables. This can look something like a table, where you have structured goals and dates by which you plan to achieve them.

An example (with some important dates filled in):

Meet with Andy (Proposal Review)	November 12 - November 15
Milestone 1 (you define)	November ??
Milestone 2 (you define)	November ??
...	...
System check with toy data	November 25
System check and prelim evaluation with real data	December 6
Final Evaluation	December 12
Final Presentation	December 16

5 Team Structure

Finally, list each of your team members as well as a (rough) breakdown of which team member will be handling implementation of what parts of the project.