EECS 6339
Introduction to Computational Linguistics Instructor: Nick Cercone
Winter Semester, 2015

In Class Project Presentation

Ross Kitsis – Appliocations of the n-gram model

Students make one 45-75 minute project presentation in class during the semester. The presentation will inform the class of the student's project and is intended to solicit feedback from students and the instructor on the feasibility of the project, and for critical commentary to take place between presenter and class.

The presentations will be graded according to the following criteria:

1. Organization -20%	20
2. Clarity -20%	19
3. Presentation-25%	
• competence-15%	14
• content-10%	10
4. Understanding-25%	24
5. Expected Outcomes-10%	10
6. Total	**97**

Please try to keep within the times allotted for the presentations and make them crisp and to the point. Don't read your slides or make them too "busy". Try to get across the ideas and not get bogged down into minute details. Practice your presentations with friends first. Good luck.

Comments: start time: 10.09; end time: 11.05
Outline (n-grams, n-gram model, applications); n-grams (unigrams, bigrams, trigrams); n-gram model (reference to Claude Shannon) – a statistical language model; definition to model; problem: sparse data illustrated with the Shakespeare example; Zipf's law in natural language processing; solution of sparse data – smoothing (3 types – additive, interpolation, backup); evaluating n-gram models – 2 ways (intrinsic and extrinsic); Applications – 6 examples (spelling correction – 3 ways); author attribution; malicious code detection; SPAM detection; textprediction; and sentiment analysis; conclusions

Overheads were well designed, elegant and well presented. Presentation was expressed well but, at times, a little rushed. Ross showed good communications skills.