

Comparison of classifiers for speech act tagging techniques

Team members: Atul Mohan <atumohan@iu.edu>
Chitesh Tewani <ctewani@iu.edu>
Prashanth Balasubramani <prasbala@iu.edu>

Abstract:

This project aims at drawing a comparison among the various speech act tagging techniques. A speech act is an utterance that serves a function in communication. Speech act is categorized into everyday conversations in form of an apology, greeting, request, conversation etc. We intend to perform the comparison by developing classifiers based on the following publications:

1. Stolcke, Andreas, Ries, Klaus, Coccaro, Noah, Shriberg, Elizabeth, Bates, Rebecca, Jurafsky, Daniel, Taylor, Paul, Martin, Rachel, Meteer, Marie, and Van Ess-Dykema, Carol. 2000. Dialogue act modeling for automatic tagging and recognition of conversational speech. *Computational Linguistics* 26(3): 339-371.
2. Ritter, A., Cherry, C., & Dolan, B. (2010). Unsupervised modeling of twitter conversations.
3. Zhang, R., Gao, D., & Li, W. (2012, April). Towards scalable speech act recognition in Twitter: tackling insufficient training data. In *Proceedings of the Workshop on Semantic Analysis in Social Media* (pp. 18-27). Association for Computational Linguistics.
4. Jeong, M., Lin, C. Y., & Lee, G. G. (2009, August). Semi-supervised speech act recognition in emails and forums. In *Proceedings of the 2009 Conference on Empirical Methods in Natural Language Processing: Volume 3-Volume 3* (pp. 1250-1259). Association for Computational Linguistics.

We would model each of these classifiers and test on large corpus of data comprising of twitter conversations. We will cumulate the findings of the results from the classifiers. We intend to obtain the training data set for each of these papers which can be used for the classifiers that would be built.