

SAMPLE: Academic Copyediting and Proofreading

Subject: information technology academic research, Chicago Manual of Style editing

Douglas Cramer: douglasramer.com, douglasramer@live.com

Excerpt: Line Edit of Technology and Linguistic Theory Paper

2.2.2 Information Extraction (IE)

IE systems tend to differ from traditional NLP_[CRM1] systems by a tendency to focusing on specific instead of general tasks, and using efficient scalable methods with a reduced linguistic theory. The best information extraction is demonstrated in the Message Understanding Conferences (MUC) sponsored by the U. S. government [52, 53, and 54] demonstrate high-quality information extraction. Its inclusion~~The importance of this field all over the world is prove by being~~ in the NIST and DARPA_[CRM2]-sponsored Text Retrieval Conferences (TREC) [97, 98] which are having the NIST and DARPA sponsorship improves the global importance of this field. IE tasks include text classification and s/categorizations; document summarization; and the design of domain-specific search engines, document summarizations, factual databases, and front ends for Kknowledge Ddiscovery in Ddatabases (KDD). While . . . Most IE works were done so far~~IE has predominately been used~~ in the area of text classification and categorization, but IE systems are just now showing potential for application in knowledge discovery in databases (KDD). Three biggest key approaches to IE are statistical analysis, inductive learning, and computational linguistics. S, however, some systems incorporate a combination of strategies to get better performance.

Text Classification and Categorization

A~~In past,~~ automatic text classification and categorization was have historically been approached in many ways, but always with the goal of less-reducing time-consuming human input or the “knowledge acquisition bottleneck.” McCallum et al. work on text classification by bootstrapping with EM_[CRM3] and shrinkage. [48]. I~~Interestingly,~~ these authors use no labeled documents, and instead using a set of keywords instead for the purpose to generate preliminary labels first. These are, which will be used alongside with a pre-defined hierarchy of classes as input forte a bootstrapping algorithm that will produces a naive Bayes classifier. Since there are so many parameters, the statistical technique shrinkage is added to the bootstrapping algorithm to deal with sparse data. For some text categorization problems, Cohen and Singer applied machine learning algorithms in accordance with context-sensitive way methods for some text categorization problems. The context-sensitive methods – the set-valued relational learner RIPPER