**STATS 517 Lit Review**

**Reviewer:** Dalyn McCauley

**Reviewee:** Yuha Yi

**Paper:** Machine Learning in Automated Text Categorizatio

**Length of Paper:**

Yuha’s review of his paper was an excellent read. It was clear, concise and most importantly understandable to a laymen like me. The paper was an appropriate length for the breadth of concepts covered.

**Depth of Understanding:**

It is clear from both Yuha’s presentation and his paper that he has a great depth of understanding on the topic of text mining and document categorization. Yuha’s use of examples show that he understands both the methodology and the application of text categorization models. The paper is written such that I can feel the authors interest and excitement about the topic.

**Completeness:**

The review was complete, and touched on all of the major points of the paper. Yuha did a great job of explaining the step by step methodologies of text mining, and also comparing the various approaches. He even gave examples of how different document structures can impact the machine learning algorithms. After thoroughly explaining the method, Yuha explains the metrics of success when assessing the accuracy of a model. He also explored the issues facing the technology, as well as where the field is going next.

**Difficulty:**

The models presented in the paper seem to be sufficiently difficult. They utilize many aspects of machine learning and text mining, and as Yuha stated, it is highly technical in nature. Despite this, the review was an easy read.

**Comprehensive:**

The paper was a comprehensive literature review. Yuha made sure to define all critical terms so the reader had a full understanding of the methods, as well as give real world examples when appropriate. He showed a few figures that complimented his points and made a well-rounded paper.

**Clarity:**

The paper was very clear. It was written to give the reader a broad understanding of machine learning in text categorization, while not getting into confusing technical details. Any student, even without a statistical or mathematic background, could read this paper and walk away with a full understanding of the methods, issues, and implications of text categorization.