**Beyond the Stars: Improving Rating Predictions using**

**Review Text Content**

**1. INTRODUCTION**

A typical application over reviews is recommendation systems, in which users do not search reviews directly but are suggested products that would best match some definition of preference.

The goal of the **URSA** (User Review Structure Analysis) project is to provide a better understanding of user reviewing patterns and to develop tools to better search, understand and access user reviews.

**2. STRUCTURE IDENTIFICATION AND ANALYSIS**

**2.1 Data Set**

**2.2 Text Review Classification**

As the first step of our project, we analyzed the data to identify categories specific to the restaurant reviews domain.

*2.2.1 Manual Sentence Annotation*

*2.2.2 Automatic Sentence Classification*

**2.3 Text Review Analysis**

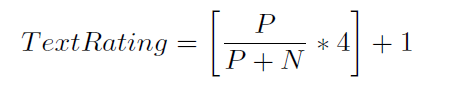
*2.3.1 User Reviewing Trends*

*2.3.2 Comparing Star Rating with Sentiment*

**3. RATING PREDICTION**

**3.1 Evaluation Setting**

**3.2 Sentiment Based Text Rating**



*3.2.1 Average Review-based Prediction*

*3.2.2 Average Metadata-based Prediction*

*3.2.3 Average Topic-based Prediction*

**3.3 Regression-based Text Rating**

*3.3.1 Regression Method*

Regression allows us to learn weights to be associated with each sentence type to account for varying importance of the sentence topics.

*3.3.2 Two-Sentiment Regression*

**4. PERSONALIZED RECOMMENDATIONS**

**5. RELATEDWORK**

**6. CONCLUSIONS AND FUTUREWORK**

In this paper, we presented the user reviews classification and analysis effort performed as part of our **URSA** project. Our main contribution is the assessment of the impact of text-derived information in predicting the rating of a review in a recommendation system. We show that both topic and sentiment information at the sentence level are useful information to leverage in a review.