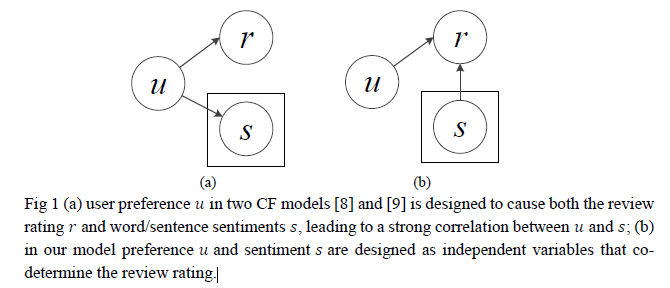
**Unifying Topic, Sentiment & Preference**

**in an HDP-Based Rating Regression Model for Online Reviews**

**1 Introduction**

In this paper we build a new review model Topic-Sentiment-Preference Regression Analysis (TSPRA) based on the HDP framework. In this paper we develop a new rating regression model to compute word sentiments automatically based on online reviews and their ratings.



**2 Related Work**

Two main approaches are studied in recent publications regarding online review rating regression: the *collaborative filtering* (CF) approaches and the *sentiment* approaches.

Many of the above-mentioned works are developed from the topic model framework, typically LDA [4], due to the fact that a reviewer usually comments on different aspects of a target item and the strong analogy between product aspects and topics.

**3 Model Construction**

**3.1 Model Specification**

Our model consists of four parts: the topic part, the sentiment part, the user preference part and the rating regression part. These four parts together describe the generative process of observable data: review text and ratings

**3.2 Inference**

**3.3 Prediction**

A regression model’s performance is commonly evaluated by its ability to predict, so is

our regression model.

**4 Experiments & Evaluations**

**4.1 Prediction Performance Evaluation**

Absolute error – the absolute value of true rating minus prediction, which is a very straight-forward measure.

Pearson’s correlation – measures how well the predicted ratings correlate with their corresponding true ratings.

Number of inverted pairs – counts the actual number of mis-ordered pairs of predicted ratings

**4.2 Sentiment Analysis**

In this section we compare the our model’s sentiment polarity evaluations to SenticNet.

**4.3 Critical Aspects**

Due to the decoupling of user preference from sentiment, we find another practical use of our model in addition to sentiment analysis -- identifying those product aspects with high user preference but ow sentiments (aspect sentiment is derived from word sentiments). The results can help business decision makers to quickly recognize certain important product aspects they should pay attention to, which we call “*critical aspects*”.

**4.4 Experiments on Parameters**

The neutral rating 𝜇 and rating noise 𝜎 are two main parameters introduced by our model TSPRA.

**5 Conclusion**

We have proposed a model that considers topics, sentiments and user preferences as integrated factors to explain review texts and ratings.