**Mining and Summarizing Customer Reviews**

**1. INTRODUCTION**

As indicated above, our task is performed in three main steps:

(1) Mining product features that have been commented on by customers.

(2) Identifying opinion sentences in each review and deciding whether each opinion sentence is positive or negative.

(3) Summarizing the results.

**2. RELATED WORK**

**2.1 Subjective Genre Classification**

**2.2 Sentiment Classification**

**2.3 Text Summarization**

Existing text summarization techniques mainly fall in one of the two categories: template instantiation and passage extraction.

**2.4 Terminology Finding**

In terminology finding, there are basically two techniques for discovering terms in corpora: symbolic approaches that rely on syntactic description of terms, namely noun phrases, and statistical approaches that exploit the fact that the words composing a term tend to be found close to each other and reoccurring.

**3. THE PROPOSED TECHNIQUES**

Note that POS tagging is the part-of-speech tagging from natural language processing, which helps us to find opinion features.

**3.1 Part-of-Speech Tagging (POS)**

Product features are usually nouns or noun phrases in review sentences.

**3.2 Frequent Features Identification**

For this purpose, we use association mining to find all frequent item sets. In our context, an itemset is simply a set of words or a phrase that occurs together in some sentences.

**Compactness pruning**: This method checks features that contain at least two words, which we call *feature phrases*, and remove those that are likely to be meaningless.

**Redundancy pruning:** In this step, we focus on removing redundant features that contain single words.

**3.3 Opinion Words Extraction**

**Definition:** *opinion sentence*

If a sentence contains one or more product features and one or more opinion words, then the sentence is called an *opinion* *sentence.*

**3.4 Orientation Identification for Opinion Words**

**3.5 Infrequent Feature Identification**

We use the *nearest* noun/noun phrase as the noun/noun phrase that the opinion word modifies because that is what happens most of the time.

**3.6 Predicting the Orientations of Opinion Sentences**

**3.7 Summary Generation**

For each discovered feature, related opinion sentences are put into positive and negative categories according to the opinion sentences’ orientations. A count is computed to show how many reviews give positive/negative opinions to the feature.

All features are ranked according to the frequency of their appearances in the reviews.

**4. EXPERIMENTAL EVALUATION**

**5. CONCLUSIONS**