

**Complementary Aspect-based Opinion Mining**

**Abstract:**

Aspect-based opinion mining is finding elaborate opinions towards a subject such as a product or an event. With explosive growth of opinionated texts on the Web, mining aspect-level opinions has become a promising means for online public opinion analysis. In particular, the boom of various types of online media provides diverse yet complementary information, bringing unprecedented opportunities for cross media aspect-opinion mining. Along this line, we propose CAMEL, a novel topic model for complementary aspect-based opinion mining across asymmetric collections. CAMEL gains information complementarity by modeling both common and specific aspects across collections, while keeping all the corresponding opinions for contrastive study. An auto-labeling scheme called AME is also proposed to help discriminate between aspect and opinion words without elaborative human labeling, which is further enhanced by adding word embedding-based similarity as a new feature. Moreover, CAMEL-DP, a nonparametric alternative to CAMEL is also proposed based on coupled Dirichlet Processes. Extensive experiments on real-world multi-collection reviews data demonstrate the superiority of our methods to competitive baselines. This is particularly true when the information shared by different collections becomes seriously fragmented. Finally, a case study on the public event “2014 Shanghai Stampede” demonstrates the practical value of CAMEL for real-world applications.

**Existing System:**

Indeed in the literature, there have been some excellent works on cross-collection topic modeling. However, they either pay less attention to the complementarity aspects across collections, or focus solely on topics and aspects without considering the opinions. Therefore, further study is still in great need for building a cross-collection aspect-based opinion mining model, based on which the diversity and complementarity in both aspect and opinion could be learned across collections containing substantially asymmetric information, e.g., the news collection with clear aspects versus the tweet collection with strong opinions.

**Proposed System:**

We propose CAMEL (Cross-collection Auto-labeled MaxEnt- LDA), a novel topic model for complementary aspect-based opinion mining across asymmetric collections. To our best knowledge, our work is among the earliest studies in this direction. CAMEL is essentially a type of cross-collection LDA model, which models aspect-level opinions and gains information complementarity by learning both common and specific aspects across different collections. By keeping all the corresponding opinions for both common and specific aspects, CAMEL is also capable of conducting contrastive opinion analysis. Moreover, to boost CAMEL, we propose AME, an automatic labeling scheme for maximum entropy model, to discriminate aspect and opinion words without heavy human labeling. It is further enhanced to the so-called EAME scheme by employing the word embedding-based similarity. Finally, we propose CAMEL-DP, a nonparametric alternative to CAMEL.