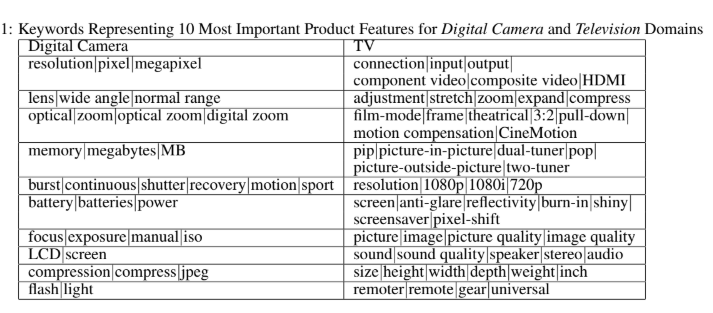
Voice of the Customers: Mining Online Customer Reviews for Product

Feature-based Ranking

인트로:

1. 제품군 별 제품 features를 가져온다
2. 문장을 features별로 나눈다.
3. 문장들을 4개로 구분한다; 긍정주관, 부정 주관, 긍정 비교, 부정 비교

Indentifying product features



1. 문장에 동의어/동의문장이 있으면 ‘describing feature’라고 구분한다.

Sentence labeling

1. MxTerminator[11] to split reviews into sentences because a typical customer review comprises of several sentences.

Identifying comparative sentences

1. 문장 내 단어들 중 KeyWords 사전에 존재하는지 확인
2. 단어들의 품사 태그 중 JJR, RBR, JJS, RBS를 구별한다
3. 문장 내 사전 정의된 패턴이 있는지 확인.
4. 문제점: “I bought this camera for my son because he got a higher grade in his second statistical exam.” 다른 제품과 비교하는 문장이 아님에도 higher 때문에 선택될 수 있음. 제품과 제품간의 비교 문장만 가져오기 위해 논문에서는
5. 해결: We use only comparative sentences which contain at least one product name which is different from the product the sentence is describing while building our ranking model. 비교 문장 안에 describing product 제외하고 다른 제품의 이름이 나오면 ‘비교 문장’으로 인식.

Identifying sentence sentiment orientation(극성 분류)

1. we use a simple yet powerful method by utilizing a positive word set(POS) and a negative word set(NEG) developed in the MQPA project[19]. 이 외에 필요 단어들을 커스텀 사전으로 구축.