**< reviews로 알아낸 제품별 특성으로 제품 순위 매기기>**

1. 소비자가 중요하게 보는 제품 특성 결정

2. subjective, comparative 문장으로 구분

Table 1는 제품 특성별 synonym sets. 고객들이 다 다르게 표현한 제품 특성을 우리가 정한 feature 열 개 중 하나로 분류함. To test our feature-finding 모델, 리뷰 문장 1000개 랜덤 추출 후 제품 특성 라벨링 함. The precision and the recall of the keyword 디카: 0.853 TV: 0.807. 리뷰 데이터셋 1516001 문장 중 16%만 제품 특성 관련임. 각 문장에 제품 특성 태그하기: 문장이 synonym set의 word나 phrase 하나라도 포함하면 그 제품 특성으로 분류함.

**[camera]**

**flash** indoor and low-light photography.

**Battery** sought-after feature that details the kind of batteries used.

**Focus** auto-focus or manual focus capabilities.

**Lens** professional photographers purchasing high-end cameras.

**Optical** encompasses digital zoom and optical zoom.

**LCD** the digital display/screen that lets a user see how a photo will look like.

**Resolution** sharpness , or detail, of a picture.

**Burs**t rapid fire and continuous shooting capabilities.

**Memory** the number of pictures that can be taken and

**Compression** how file size of a photo is shrunk.

**[TV]**

**the Sound feature** for users interested in audio quality (some TVs come with an extra set of speakers to create surround sound).

**Reflectivity/Anti-glare** for the viewing experience.

**Size** size,height, weight of a television screen.

**Connection** the number and type of input ports available for hooking up devices to the television.

**Picture quality** richness/quality of the images displayed

**Users** remote control device available with the television.

**Resolution** number of pixels or lines displayed on the screen.

A**djustment** ability/mode that expands or compresses an image to fill the screen better.

**Picture-in-picture(PIP)** to watch two channels at once.

**Film-Mode/CineMotion** the movie-watching experience

리뷰를 문장 단위로 나누기 위해 MxTerminator[11] 사용

**[리뷰 문장 두 종류]**

**Subjective Sentence(SS)** : 제품 관련 direct praise or deprecation

Ex. This camera has excellent shutter speed.

**Comparative Sentence(CS)** indirectly express an opinion by comparing two products.

Ex. I think the coolpix is better than the canon sd1200.

**Product Comparative Sentence(PCS)** comparative sentence이면서 at least one product name 포함하는 문장. Ex. This TV has much better sound quality when compared to the sony bravia.

**[CS 중 PCS 찾기**]

Use keyword comparison(KW=126words, some are explicit(“outperform, exceed, compare, superior, etc.”) and others are implicit(“prefer,choose))

POSTags 위해 CRFTagger[12]사용

• 문장이 KW의 comparative keywords 하나라도 포함하는지 확인;

• Recognize any words with POS tags ∈ JJR(comparative adjective), RBR(comparative adverb), JJS(superlative adjective), RBS(superlative,adverb)

• Scan if any predefined structural patterns are present in the sentence (as <word> as, the same

as, similar to.)

**dynamic programming technique**: product-product comparison pairs in a comparative sentence. only use comparative sentences which contain at least one product name, precision of 82% and a recall of 80% approximately.

MQPA project[19]에서 **positive word set(POS)/negative word set(NEG)**.

If the sentence contains a word in positive word set, label this sentence a positive tag.

however using some negations in their reviews.=> add 28 negation words manually.

positive set 1974 words, negative set 4605 words

3. 제품 특성 노드-엣지 그림 그리기 (가중치= 상대적 퀄리티)

Gf = (V, E)

V 노드= {pi| 각 node=product, 0 < i < n},

E 엣찌 = {ek = (pi, pj ) pi--->pj | Wei=엣찌(ei) 가중치 0 < i, j < n(제품수),0 < k <m(엣찌수)}

제품Pi 리뷰 중 특성f 내용의 comparative 문장이 Pi랑 Pj를 비교하면, Pj-🡪Pi 엣찌를 그린다.

그 담에 엣찌 가중치(PC갯수/NC갯수)구한다. (PC=긍정comparative sentences)

Pi>Pj=PC(Pi, Pj )= Pi가 낫다/Pi<Pj =NC(Pi, Pj )=Pj가 낫다

그 담에 노드Pi 가중치(제품퀄리티)= PS갯수/NS갯수 (PS=긍정subjective sentences)

4. 페이지랭크로 순위 매기기

5. 아마존닷컴의 디카, TV 카테고리 리뷰(<https://www.amazon.com/s?i=photo&bbn=502394&rh=n%3A172282%2Cn%3A%21493964%2Cn%3A502394%2Cn%3A281052%2Cp_n_shipping_option-bin%3A3242350011&s=review-rank&dc&fst=as%3Aoff&pf_rd_i=16225009011&pf_rd_m=ATVPDKIKX0DER&pf_rd_p=82d03e2f-30e3-48bf-a811-d3d2a6628949&pf_rd_r=8QQMENRD5F91P2Q66PYG&pf_rd_s=merchandised-search-4&pf_rd_t=101&qid=1564555652&qid=1564555652&rnid=502394&ref=sr_nr_n_4>)

사용해서 모델 성능 테스트하기