

인터넷 게시판과 블로그에 나타난 긍정적/부정적 상품평 분석 도구의 설계 및 구현

김동의, 이상곤

전주대학교 컴퓨터공학과

e-mail : ksh6438@naver.com, and samuel@jj.ac.kr

Design and Implementation of Positive/Negative Review for Product Evaluation Tool in Internet BBS or Blogs

Donguei Kim and Samuel Sangkon Lee

Dept of Computer Science & Engineering, Jeonju University

요 약

본 논문은 인터넷에서 사용자가 물품을 구매하고자 할 때 구매를 원하는 것의 상품평을 미리 볼 수 있는 프로그램을 설계하고 구현하였다. 구매 물건의 동종 모델에서 어느 제품을 구매하여야 할지 사용자의 선택이 어려운 경우 고객의 의사결정에 도움을 주고자 해당 상품을 이미 구매한 고객들의 의견 중에서 긍정적인 면과 부정적인 면 중 어떤 성격이 더 강한지를 판단하는 상품평 분석 도구이다. 본 프로젝트에서는 사용자의 물품 구매를 돕는 상품평 분석 도구의 기능은 첫째, 인터넷 게시판과 블로그(blog)의 글을 수집하고, 이를 분석해서 긍정/부정/보통의 의견들을 그래프로 표시하여 구매 의사 결정에 도움이 준다.

1. 서론

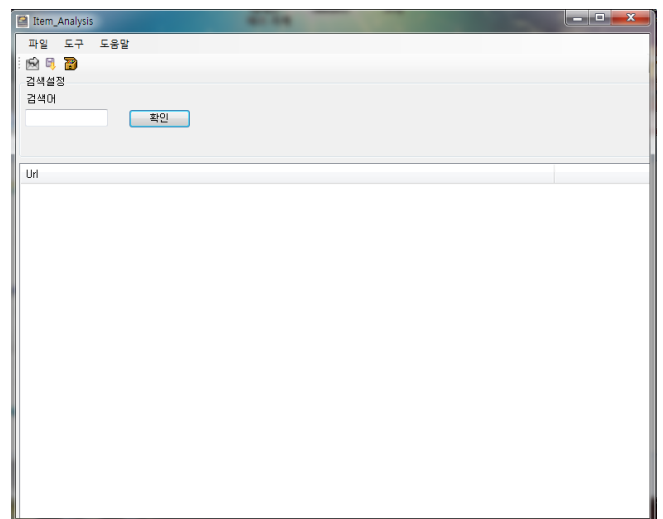
현대에는 시장이나 마트보다 인터넷 상에서 더 많은 물품의 소비가 이루어지고 있다. 사용자가 물품 구매 시 동일한 기능의 제품이 너무 많아 어떤 제품을 구매하여야 할지 그 의사결정이 어려운 경우가 있다. 이 때 동종 물품을 이미 구매하여 사용해 본 다른 사람들의 상품평을 보고 사람들의 의견 중 긍정적인 면이 많을 경우 물품을 구매하는 경향[1, 2, 3, 4]이 있다.

본 논문의 개발 목적은 상품평을 분석하는 데에 있어, 해당 상품의 긍정적인 단어와 부정적인 단어의 지식 사전(부록에 일부 단어를 제시)을 구축[9]하고, 이를 상품평과 매칭하고 구매 의사 결정에 도움을 주는 것이다. 상품평은 인터넷의 게시판이나 네이버 블로그에서 자동으로 수집한다.

본 논문에서 다루는 부분은 상품평 단어 분석이지만 처리가 이뤄지는 주요 핵심 시스템은 웹문서의 자동 수집과 상품평의 빠른 검색이 강점이다. 따라서 본 논문에서 제안하는 시스템을 이해하기에는 웹 태그의 구조 분석과 데이터베이스의 구조, 처리에 대한 이해가 필요하고, 추가적으로 매칭하는 알고리즘의 이해가 필요하다. 데이터베이스의 설계는 일단 요구 분석, 개념 설계, 논리 설계, 물리 설계, 유지보수 등이 필요하다.

2. 분석기의 기능

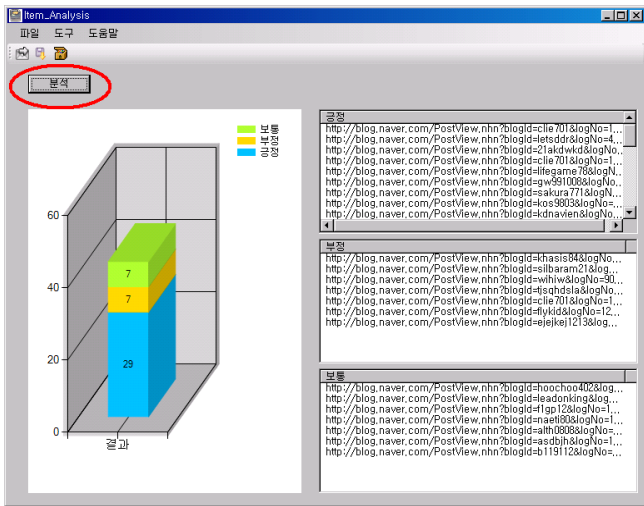
본 장에서는 검색과 태그를 분리하는 부분에 대한 기술적인 내용을 서술한다. 본 프로젝트의 구현은 Visual Studio 2008을 이용하였고, 서비스 환경으로 C#를 사용하였다. 운영 체제는 MS Windows 7 기반 하여 설계하였고, 해상도는 1,440 × 900을 기준으로 구현하였다.



(그림 1) 시스템의 기본 화면

본 논문에서는 검색 및 수집 부분은 Naver API를 이용하여 XML 결과를 받아서 태그를 분리하고, 실제 주소에 접근하여 해당 내용을 얻어, 매칭 모듈로 데이터를 넘겨주고, 이를 미리 구축된 사전과 매칭 하여 결과를 도출[7, 8, 9]하고, 그래픽컬 하게 보여준다. 본 프로그램은 네이버 블로그 검색을 통하여 블로그에 등록된 상품평을 수집하고 이를 분석하여 사용자에서 도움을 준다.

위의 (그림 1)은 본 프로젝트에서 개발한 프로그램의 주요 화면이다. 검색어를 넣고 확인 버튼을 누르면 아래의 리스트 뷰에 URL 결과를 표시해 준다. 사용자는 해당 결과를 더블 클릭할 경우, 해당 URL에 곧바로 접근할 수 있다. 검색어에 검색할 물품명을 넣고 검색하면 리스트 뷰에 주소가 누적된다. 상자를 여는 모양 아이콘을 누르면 인터넷 속도에 영향을 받지만 상품평의 분석을 시작합니다.



(그림 2) 분석 결과 화면

위의 (그림 2)와 같이 분석 버튼 위에 있는 상자를 누르면 위의 분석 결과가 화면에 출력되고, 결과 그래프와 함께 긍정/부정/보통의 리스트를 채워 넣습니다[10, 11]. 해당 링크를 누르면 해당 URL로 이동하여 상품에 대한 평을 실시간으로 확인할 수 있도록 하였다.

3. 결론

본 논문에서는 휴대폰이나 기타 상품들 관련 리뷰이나 기사들을 수집하여 분석하고 특정 상품에 대한 좋은 평과 나쁜 평[12], 그리고 보통의 평가들을 수집 분류하고 그것을 비주얼적으로 보여 주는 프로그램을 개발하였다. 향후 과제로는 스마트폰 어플리케이션으로 확장하고 분석 결과를 계속하여 DB에 축적하고 웹 서비스를 통해 다양한 정보를 추출해 낼 수 있도록 할 생각이다.

사람들이 물건을 살 때 다른 사람들의 의견이나 제품의 장단점을 알고 싶어 할 것으로 생각되었고 해당 상품의 관련 댓글이나 기사를 수집해서 분석하고 이를 긍정과 부

정, 보통 의견으로 분류하여 막대그래프의 형태로 표현하고 오른쪽에는 관련 기사나 댓글을 보여준다.

본 논문의 도구는 C#을 이용하여 구현되었으며, 개발자들이 쉽게 사용할 수 있는 MY-SQL을 이용하여 데이터베이스를 설계하고 프로그램과 연동을 하도록 하였다. 인터페이스는 Naver에서 제공하는 API를 이용하여 블로그 검색 및 검색 결과를 도출한다.

참고 문헌

- [1] 강한훈 외, "k-Structure를 이용한 한국어 상품평 단어 자동 추출 방법," 정보과학회논문지 : 소프트웨어 및 응용, 제37권, 제6호, pp.470-479, 2010.
- [2] 김세종 외, "단어의 문맥적 위치와 문장 유사도를 이용한 상품 특성 추출 및 계층화," 2010 한국 컴퓨터종합학술대회 논문집, 제37권, 제1(A)호, pp.96-97, 2010.
- [3] 송중석 외, "상품평 분류 성능 향상을 위한 긍정/부정 사전 자동 구축," 2010 한국 컴퓨터종합학술대회 논문집, 제37권, 제1(A)호, pp.136-137, 2010.
- [4] 김정호 외, "어미변화를 고려한 감성 구문 패턴을 이용한 상품평 의견 분류," 2010 한국 컴퓨터종합학술대회 논문집, 제37권, 제1(C)호, pp.285-290, 2010.
- [5] 광동민 외, "의견어 중심의 의존 트리 패턴 자질을 이용한 기계 학습 기반 한국어 블로그 문서 의견 분류 시스템," 한국 정보처리학회 추계 학술대회 논문집, 제16권, 제2호, pp.321-322, 2009.
- [6] 김민석 외, "상품평을 활용한 상품 특징 기반 추천 기법," 한국 정보과학회 가을 학술대회 논문집, 제36권, 제2(A)호, pp.34-35, 2009.
- [7] 김한훈 외, "상품 리뷰의 속성별 긍정/부정 분류 시스템의 설계 및 구현," 한국 정보과학회 가을 학술대회 논문집, 제36권, 제2(A)호, pp.1-6, 2009.
- [8] 김진욱 외, "한글 텍스트의 오피니언 분류 자동화 기법," 한국 컴퓨터종합학술대회 논문집, 제36권, 제1(A)호, pp.86-90, 2009.
- [9] 육상근, "사전 기반의 한국어 상품 리뷰 의견 표현 자질 추출 및 분류 시스템," 제 30회 한국 정보처리학회 추계 학술발표 논문집, 제15호, 제2호, pp.631-634, 2008.
- [10] 배민영 외, "Topic Signature를 이용한 댓글 분류 시스템," 정보과학회 논문지 : 소프트웨어 및 응용, 제35권, 제12호, pp.774-779, 2008.
- [11] 배민영 외, "Topic Signature를 이용한 댓글 분류 시스템," 한국 컴퓨터 종합 학술대회 논문집, 제35권, 제1(A)호, pp.80-81, 2008.
- [12] 송호진 외, "One-class 문서 분류를 위한 긍정 자질과 부정 자질의 결합," 제17회 한글 및 한국어 정보처리 학술대회, pp.35-42, 2005.

[부록 A] 부정 단어의 예

가슴이 터지다	가자미눈을 뜨다	가혹	가혹하다	불만이 가득하다	불쾌	불쾌스러운
가혹한	개같은 놈			불쾌하다	불쾌한	불편함이 있다
개념을 밥말아 드셨군요		개새끼		비참하다	빡세다	빡통
거품을 내다	거품을 폼다	검토해		뺨이 나다	빼꾸	살을 떨다
결장이 터지다	경멸	경멸하다		삼질하다	새끼	소름끼치는
경멸한	계집애	꼴때리다		소름끼치다	속니를 갈다	속을 뒤집다
패섬하다	패섬한	구라치다		속이 굵다	속이 부글부글 피다	
구리다	구역질	구역질나는		속이 치밀다	속이 타다	속이 터지다
구역질나다	그 판식	그것밖에 못하니		송곳니가 방석니가 되다		숨막히는
그닥	깔삼하다	갑치지마		싫다	싫어하다	심사가 뒤집히다
깨지다	꺼져	꼬라지		싸가지 없는	싸잡다	싸잡아
꼬봉	꿀값	꿀사나운		싸잡아서	쌍까다	쓰레기 같은 놈
꿀사납다	꿀통	끓어오르는		쌈쓸하다	쌈쓸하다	쌈쓸한
끓어오르다	나랑은 상관없다	나발불다		씨발	씨발새끼	씹다
나쁘다	나쁜	나쁜 년		아가리	아랫것들	아쉬운
나쁜 놈	너석	노하다		아쉽다	악을 쓰다	야려보다
노한	놈	눈갈		야리까다	약오르는	약오르다
눈알이 빠지다	눈알이 시뻘게지다			억울하다	억울한	언짢다
눈에 날이 서다	눈에 모를 세우다			언짢은	얼굴이 붉으락푸르락해지다	
눈에 불을 켜다	눈에 불이나다			얼굴이 새파래지다	얼굴이 파랗게 질리다	
눈에 쌍불을 켜다	눈에 쌍심지가 돌다			얼굴이 파래지다	역겨운	역겹다
눈에 쌍심지를 세우다	눈에 쌍심지를 켜다			열등감	열을 받다	
눈에 천불이 나다	눈에 칼날이 서다			열이 꼭두까지 오르다	열이 뻗치다	
눈을 까뒤집다	눈을 치뜨다	눈을 흡뜨다		열이 상투 끝까지 오르다	열이 오르다	
눈이 곤두서다	눈이 돌아가다	눈이 뒤집히다		열이 터지다	열이나다	
눈이 시퍼래지다	단점	닭대가리	대가리	열통이 터지다	오장이 뒤집히다	왕재수
도끼눈을 뜨다	도야지 새끼 같은 놈			욕	울화가 터지다	울화통이 터지다
독사눈을 뜨다	돼져버려	돼지 새끼 같은 놈		위협	이 녀석	이가 갈리다
두 눈을 부릅뜨다	따위	뺨다		이빨까다	이새끼	이씨발새끼
뺨은	또라이	뜯만이		이자식	입닥쳐	입에 거품을 물다
똥	띠겁다	머저리		저 퀄리티	절망	절망스러운
멍청이	모욕	모욕적		절망한	존나	줄라
모욕적이다	몸을 떨다	무서운		주둥아리	주먹을 떨다	주먹을 불끈 쥐다
무섭다	문제가 발생했다	문제도 발생했다		죽는다	죽어	죽어라
미운	미워 죽겠다	미친년		죽여버린다	취 패버린다	증오
미친놈	밧다	바보		증오하다	증오한	지랄하네
방망이로 맞고 홍두깨로 때린다				진상	짓	짜증
방망이로 얻어맞은 놈 홍두깨로 친다				짜증난다	짜증스러운	조개다
배반	배반감	배반하다		쪽박차다	쫄다구	찌그러뜨린다
배신	배신감	배신당했다		참답하다	치를 떨다	칼을 갈다
배신하다	머그가 많다	머그폰 이다		칼을 폼다	필필 뛰다	폭탄
별로다	병신	복수심		피가 거꾸로 솟다	피가 굵다	
복수하다	복장을 뒤집다	복장이 타다		피를 토하다	핏대를 돋구다	핏대를 세우다
복장이 터지다	붕의 눈을 뜨다	부레가 굵다		핏대를 올린다	하나 잘하는 게 없구나	
부아가 굵다	부아가 치밀다	부아통이 터지다		한뼉히다	호구	화나다
북받친	분개한	분노	분통이 터지다	화났다	화통이 터지다	환불
불덩이가 치솟다	불량	불만	불만과 분노			환불요구
불만스러운	불만스럽다	불만은 폭발				

[부록 B] 긍정 단어의 예

가슴에 와닿다	가슴을 울리다	싱그럽다	쓸개가 퍼지다	아름다운
가슴을 적시다	가슴을 파고들다	아름답다	아주 멋져요	아주 멋진
가슴을 펴다	가슴이 떨리다	양이 차다		어깨가 가볍다
가슴이 뜨거워지다	가슴이 몽클하다	어깨가 들썩들썩		어깨가 올라가다
가슴이 벅차오르다	가슴이 부풀다	어깨가 으쓱거리다		어깨가 으쓱해지다
가슴이 뿌듯하다	가슴이 시원하다	어깨를 펴다		어깨춤을 추다
가슴이 울렁거리다	가슴이 쩡하다	어깨춤이 절로 나다		얼굴이 퍼지다
가슴이 트이다	가슴이 후련하다	예쁘다	예쁜	우월
가슴이 훈훈해진다	감격	우월하다	우월한 듯	웃다
감격스럽다	감격스러운	웃음	월등	유쾌하다
감동하다	감동	유쾌한	입시울이	함박꽃 같이 벌어지다
감사하다	감사하다	입을 다물지 못하다		입이 찢어지다
감사한	경쾌하다	입이 함박만하다		잘 할 수 있어요
고동이 치다	고마운	잘됐네	잘됐어	잘한
고무적인	고맙게	잘할 수 있어	잘했다	잘했어
기분 좋은	기대	재미	재밌	좋다
기쁘다	기쁠	좋아	좋은	좋더군요
꽃	날아갈 듯한	즐겁다	즐거운	즐겁게
눈물이 나다	넌 최고야	즐겁다	추천	축하하다
눈시울이 뜨거워지다	넌 최고야	축하한다	축하	축하하다
눈시울이 시큰해지다	넌 최고야	콧날이 찡하다	콧날이 시큰해지다	
다리를 뻗고 자다	넌 최고야	콧날이 쩡하다	콧노래가 나오다	
다정	넌 최고야	콧노래를 부르다	쾌활하다	
다정하다	넌 최고야	쾌활한	탐난다	통쾌하다
대단	넌 최고야	통쾌한	파이팅	편의를 더해주는
대단하다	넌 최고야	포근하다	포근한	풍부한
등따습고 배부르다	넌 최고야	합격	합격하는	합격하다
마음에 쏙드는	넌 최고야	합격한	행복	행복하다
만족	넌 최고야	허리가 끊어지다		허리가 부러지다
만족하다	넌 최고야	허리를 잡다		호감가는
명랑하다	넌 최고야	호감가다	홀가분하다	홀가분한
믿는다	넌 최고야	환상적이다	환상적인	환하다
반갑다	넌 최고야	환한	활기있는	활기있다
밝다	넌 최고야	활발한	휘파람을 불다	흐뭇하다
배꼽을 빼다	넌 최고야	흐뭇한	흡족하다	흡족한
배꼽이 빠지다	넌 최고야	희망	희망적인	희망차다
벅차다	넌 최고야	희망찬		
뿌듯	넌 최고야			
사랑스러운	넌 최고야			
산뜻하다	넌 최고야			
살맛나다	넌 최고야			
상쾌하다	넌 최고야			
상큼한	넌 최고야			
속이 시원하다	넌 최고야			
손뼉을 치다	넌 최고야			
순수하다	넌 최고야			
시원한	넌 최고야			
신뢰	넌 최고야			
신명이 나다	넌 최고야			
신용	넌 최고야			
십년 묵은 체증이 내려가다	넌 최고야			