

감정 분석 기반의 사용자 피드백을 이용한 클라우드 서비스 평가 기법[†]

(Cloud Service Evaluation Techniques Using User Feedback based on Sentiment Analysis)

윤 동 규 [‡] 김 응 수 [§] 박 준 석 [§] 염 근 혁 [¶]
 (Donggyu Yun) (Ungsoo Kim) (Joonseok Park) (Keunhyuk Yeom)

요 약 클라우드 컴퓨팅이 IT 업계의 화두로 부상하면서 다양한 유형의 클라우드 서비스들이 나타났고, 다수의 클라우드 서비스들 중 사용자가 원하는 서비스를 선택하는 과정의 복잡성을 해소하기 위해 클라우드 서비스 브로커 (Cloud Service Broker, CSB) 기술이 등장하였다. CSB의 핵심적인 기능 중 하나는 사용자에게 최적의 클라우드 서비스를 추천해주는 것이다. 일반적으로 CSB에서 클라우드 서비스 추천을 위해 서비스 사용자들로부터 서비스에 대한 평점을 피드백으로 받아 서비스를 평가하는 방법을 사용할 수 있다. 그러나 사용자마다 평점을 매기는 기준이 다양하므로 평점만으로 서비스를 평가하기에는 신뢰도가 떨어진다. 본 논문에서는 클라우드 서비스 사용자의 리뷰에 기계 학습 기반의 감정 분석 (Sentiment Analysis) 기법을 적용하여 평점 기반 서비스 평가를 보완하는 방법과 이를 적용하여 구현한 CSB의 프로토타입을 제시한다. 또한 실제 클라우드 서비스 리뷰를 학습 데이터로 사용한 실험을 통해 감정 분석에 사용될 수 있는 여러 학습 알고리즘의 성능을 비교한 결과를 제시한다. 본 논문에서 제안하는 서비스 평가 기법은 기존의 평점 기반 서비스 평가의 단점을 보완하며 사용자 경험 측면의 서비스 품질을 반영할 수 있다.

키워드 : 클라우드 서비스 브로커, 클라우드 서비스 평가, 사용자 피드백, 감정 분석

Abstract As cloud computing has emerged as a hot trend in the IT industry, various types of cloud services have emerged. In addition, cloud service broker (CSB) technology has emerged to alleviate the complexity of the process of selecting the desired service that user wants among the various cloud services. One of the key features of the CSB is to recommend the best cloud services to users. In general, CSB can use a method to evaluate a service by receiving feedback about a service from users in order to recommend a cloud service. However, since each user has different criteria for giving a rating, there is a problem that reliability of service evaluation can be low when the rating is only used. In this paper, a method is proposed to supplement evaluation of rating based service by applying machine learning based sentiment analysis to cloud service user's review. In addition, the CSB prototype is implemented based on proposed method. Further, the results of comparing the performance of various learning algorithms is proposed that can be used for sentiment analysis through experiments using actual cloud service review as learning data. The proposed service evaluation method complements the disadvantages of the existing rating-based service evaluation and can reflect the service quality in terms of user experience.

Key words : Cloud Service Broker, Cloud Service Evaluation, User Feedback, Sentiment Analysis

1. 서 론

클라우드 컴퓨팅의 대중화와 함께 클라우드 서비스에 대한 사용자들의 선택지는 점점 더 많아지고 있고, 그만큼 사용자가 자신에게 적합한 클라우드 서비스를 찾아내는 일은 더 어려워지고 있다. 이와 같은 문제점을 해소하기 위해 클라우드 서비스 브로커 (Cloud Service Broker, CSB) 라는 개념이 등장하였다. CSB는 클라우드

서비스 제공자와 사용자 사이에서 클라우드 서비스를 추천해 주며 계약을 중개해 주는 역할을 수행한다[1].

CSB에서 핵심적으로 고려되어야 하는 사항은 사용자의 기대를 충족시켜줄 수 있는 적합한 클라우드 서비스를 추천해 주는 것이다. 따라서 높은 품질의 CSB를 구축하기 위해서는 클라우드 서비스 추천의 성능을 높여 사용자의 요구에 적합한 클라우드 서비스를 추천해줄 수 있어야 하며, 이를 위해서는 클라우드 서비스의 어떤 요소를 기준으로 적합성을 판별할 것인지 결정해야 한다. CSB는 클라우드 서비스 제공자로부터 제공받은 정보나 자체적으로 모니터링한 정보 등을 이용하여 서비스를 평

[†] 이 논문은 2016년도 정부(교육부)의 재원으로 한국연구재단의 지원을 받아 수행된 기초연구사업임 (No. NRF-2016R1D1A1B03935869)

[‡] 교신저자

가하는 메커니즘을 가질 수 있다. 이를 클라우드 서비스 평가라 지칭한다.

본 논문에서는 CSB에서 클라우드 서비스를 평가하기 위한 방법으로 클라우드 서비스 사용자의 피드백을 기반으로 서비스를 평가하는 기법을 제안한다. 이 기법은 기존의 대표적인 피드백 기반 서비스 평가 방법인 평점 기반 서비스 평가의 단점을 보완하기 위해 감정 분석 (또는 감성 분석, Sentiment Analysis) 을 적용하여 서비스를 평가하는 기법이다. 즉 사용자가 작성한 리뷰에서 서비스에 대한 사용자의 감정을 추출하여 이를 기반으로 평점을 보정한다. 그리고 기계 학습을 이용해 클라우드 서비스 리뷰에 대한 감정 분석 실험을 수행한 결과를 통해 다양한 기계 학습 알고리즘의 성능을 비교한 결과를 제시한다. 또한 제안하는 서비스 평가 기법을 적용하여 CSB 프로토타입을 구현한 결과를 제시한다.

본 논문의 구성은 다음과 같다. 2장에서는 사용자 피드백과 클라우드 서비스 평가에 대한 관련 연구들을 제시한다. 3장에서는 기존의 평점 기반 서비스 평가의 문제점을 제시하고, 이를 보완하기 위해 감정 분석 기반의 서비스 평가 기법을 제안한다. 4장에서는 실제 클라우드 서비스의 리뷰에 대하여 기계 학습을 이용한 감정 분석 실험을 수행한 결과를 제시한다. 5장에서는 제안하는 서비스 평가 기법이 적용된 CSB 프로토타입의 구조와 사용 사례를 제시하며, 6장에서 본 연구의 의의와 결론 및 향후 연구에 대하여 제시한다.

2. 관련 연구

사용자 피드백의 분류와 사용자 피드백을 이용한 감정 분석, 사용자 피드백을 이용한 서비스 평가 등과 관련하여 다음과 같은 연구들이 제시되었다.

J. Choi 등의 연구[2]에서는 사용자 피드백의 유형 (Types) 을 Survey, Question and Answer, Communication, Collaboration의 4가지로 정의하였다. 이에 따르면 본 논문에서 사용하는 피드백 방법은 사용자로부터 리뷰를 입력 받는 Communication 유형에 속한다.

E. Guzman 등의 연구[3]에서는 구글 플레이 스토어, 애플 앱 스토어에 올라온 애플리케이션에 대한 리뷰를 대상으로 감정 분석을 수행하였다. NLTK 라이브러리를 이용하여 리뷰 텍스트에 대한 전처리를 수행하고 어휘 사전 기반으로 점수를 부여하였으며, 애플리케이션의 기능 별로 사용자들이 긍정적으로 평가했는지 또는 부정적으로 평가했는지 예측하는 기법을 제시하였다. 단 이 기법은 어휘 사전 기반이므로 사전에 없는 신어 (新語) 나 특정 도메인에서 사용되는 단어에 대한 처리는 불가능하다는 한계점이 있다.

이태원 등의 연구[4]는 기계 학습 알고리즘인 서포트 벡터 머신 (Support Vector Machine, SVM) 을 이용하여 온라인 도서, 영화 리뷰에 대한 감정 분석을 수행한

결과를 제시하였다. SVM의 입력 변수로 형용사, 4품사에 대한 문서빈도, TF-IDF, 정보 획득량, 카이제곱 통계량을 각각 사용하여 입력 변수에 따른 분류 성능을 테스트하였다.

L. Qu 등의 연구[5]에서는 사용자 피드백과 정량적 성능 평가를 결합하여 클라우드 서비스를 추천해주는 기법을 제시하였다. 사용자 피드백 부분에서는 퍼지 이론을 적용하여 사용자의 주관적인 평가를 점수로 변환하는 과정을 거친다. 주관성과 객관성을 모두 반영할 수 있는 평가 방법이지만 사용자의 주관적 평가를 나타내는 단어 (good, fair, poor 등) 마다 점수를 계산하기 위한 weight 값을 수동으로 할당해야 하므로 다양한 감정 표현을 평가에 반영하기는 어렵다.

3. 감정 분석 기반의 사용자 피드백 평가

3.1 기존의 평점 기반 서비스 평가 방법

사용자로부터 서비스에 대한 피드백을 받는 가장 일반적인 방법으로는 평점을 입력 받는 방법이 있다. 5점 만점 또는 10점 만점의 척도를 가장 많이 사용하며, 가장 간단하게 사용자의 피드백을 평가에 반영할 수 있는 시스템으로 영화나 책에 대한 평가, 쇼핑몰에서 상품에 대한 평가, 앱 스토어에서 애플리케이션에 대한 평가 등 각종 평가 시스템에 폭넓게 이용되고 있다.

그러나 이와 같은 평점 기반 평가 시스템은 아래와 같은 문제점을 가지고 있다.

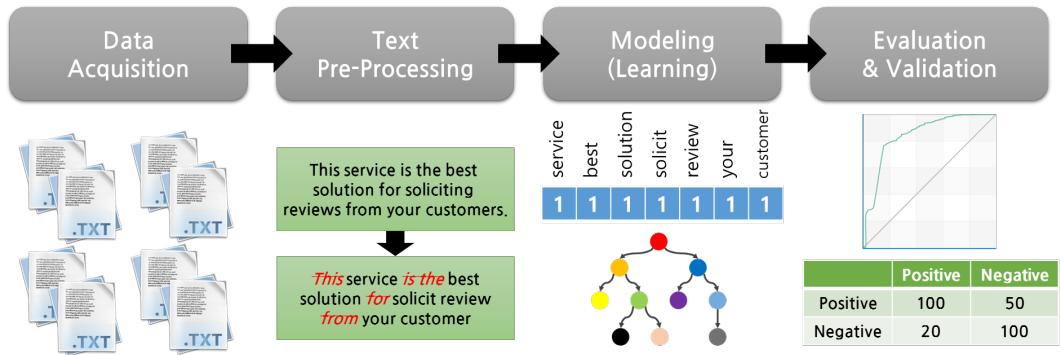
- 5단계, 10단계의 세부적인 구분이 크게 의미가 없는 경우가 많다. 예를 들어 5점 만점의 평가인 경우 대다수의 사용자들이 서비스가 만족스러웠을 경우 5점, 불만족스러웠을 경우 1점을 매기며 이 경우 2, 3, 4점의 존재는 전문가의 리뷰가 아닌 이상 크게 의미가 없다[6]. 즉 실질적으로 ‘좋아요-싫어요’ 의 이분법 평가와 크게 다를 것이 없다.

- 사용자마다 평가 기준이 다르다. 서비스의 품질이 적당한 수준만 되어도 5점 만점을 주는 사용자도 있는 반면 평가 기준이 엄격한 사용자는 좋은 서비스라는 리뷰를 남기면서도 4점을 줄 수도 있다.

본 논문에서는 이와 같은 문제를 해결하기 위한 방안으로 서비스 리뷰에 감정 분석 기법을 적용하여 서비스에 대한 사용자의 감정에 따라 평점에 보정을 가하는 방법을 제안한다.

3.2 감정 분석 프로세스

감정 분석이란 대상 (주로 텍스트) 으로부터 대상에 담겨있는 감정을 추출해내는 방법을 말하며, 텍스트 마이닝 (Text Mining) 의 하위 분야이기도 하다. 최근에는 주로 소셜 미디어에서 사용자들이 업로드하는 텍스트를 수집하여 특정 대상에 대한 여론을 조사하는 용도로 많이 사용되어 오피니언 마이닝 (Opinion Mining) 이라고도 불린다.



[그림 8] 감정 분석 프로세스

감정 분석의 일반적인 프로세스는 위 [그림 1]과 같다.

첫 번째 과정은 데이터를 수집하는 과정 (Data Acquisition) 이다. 분석 대상이 될 데이터와 학습을 위해 필요한 데이터 (Training Data) 가 필요하다. 본 연구에서는 클라우드 서비스 리뷰에 대한 감정 분석을 진행할 것이므로 클라우드 서비스 리뷰를 모은 텍스트 파일이 Training Data가 된다.

두 번째 과정은 텍스트 전처리 (Text Pre-Processing) 과정으로, 자연어 처리 (Natural Language Processing, NLP) 분야와 깊게 관련이 있는 부분이다. 첫 번째 과정에서 수집한 데이터에는 일반적으로 노이즈가 많아 그대로 학습에 사용하기에는 적절하지 않은 경우가 많다. 따라서 노이즈를 없애고 학습에 필요한 단어만을 남겨놓기 위한 전처리 과정이 필요하다. 전처리 과정에서는 문장을 단어 단위로 쪼개고 (Tokenize), 접속사, 판사, be동사 등 감정과 관계없는 단어를 삭제하고 (Remove Stopword), 단어를 원형으로 고치는 (Lemmatize) 방법 등이 사용된다.

세 번째 과정은 감정 분석 모델을 만드는 부분 (Modeling) 이다. 크게 기계 학습 기반의 방법과 사전 기반의 방법으로 나눌 수 있는데[7], 최근에는 기계 학습 기반의 방법이 많이 사용된다. 특히 기계 학습 기반의 방법은 사전에 없는 단어나 특정 도메인에서 자주 쓰이는 단어 또한 분석에 반영할 수 있기 때문에 클라우드 서비스 리뷰에 대한 감정 분석 모델을 만들기에 적합하다.

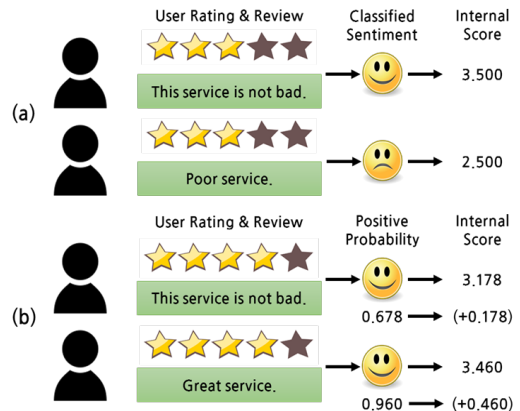
마지막 네 번째 과정은 감정 분석 모델의 성능을 검증 및 평가하는 부분 (Evaluation & Validation) 이다. 기계 학습 기반의 방법에서는 일반적으로 교차 검증 (Cross-Validation) 방법을 사용하게 되며, 평가의 척도로는 정밀도 (Accuracy), 정확도 (Precision), 재현율 (Recall), 정확도와 재현율의 조화평균인 F1 Score 등이 사용된다.

3.3 감정 분석을 적용한 사용자 피드백 평가 기법

본 논문에서 제안하는 서비스 평가 방법은 3.1절에서 제시한 기존의 평점 기반 서비스 평가에 3.2절에서 제시한 감정 분석 프로세스를 같이 적용하여 각각의 단점을

보완하는 것이다. 3.1절에서 언급하였듯 기존의 평점 기반 서비스 평가 방법에서는 ‘그럭저럭 쓸만한 서비스’ 라는 의미의 3점과 ‘그다지 추천하고 싶지는 않은 서비스’ 라는 의미의 3점을 내부적으로 동일하게 3점으로 처리한다. 그러나 평점과 함께 텍스트 형식의 리뷰를 입력 받고, 감정 분석 모델을 이용한다면 두 리뷰의 점수에 대한 차등화가 가능하다. 감정 분석 모델을 적용했을 때 전자의 리뷰는 ‘긍정적’ (positive), 후자의 리뷰는 ‘부정적’ (negative) 으로 분류되었다면 사용자가 입력한 평점은 동일한 3점일지라도 CSB 내부적으로는 각각 3.5점, 2.5점이나 4점, 2점으로 처리한다는 의미이다. [그림 2]의 (a) 는 이를 그림으로 나타낸 것이다.

위와 같이 감정 분석 모델의 분류 결과 (긍정적/부정적) 에 따라 고정적인 상수를 보정 값으로 주는 방법도 있지만, 학습 알고리즘으로 확률 알고리즘을 이용한다면 보다 세부적으로 보정 값을 부여할 수 있다. 똑같이 ‘긍정적’ 으로 분류된 텍스트라고 하더라도 60% 긍정적인 리뷰보다는 99% 긍정적인 리뷰에 더 서비스에 대해 호의적인 평가가 담겨 있을 가능성이 높다. 따라서 후자의 리뷰에 더 높은 보정 값을 부여하는 방법이 있다. [그림 2]의 (b) 는 이를 그림으로 나타낸 것이다.



[그림 9] 감정 분석 결과를 적용한 평점 보정

4. 클라우드 서비스 리뷰에 대한 감정 분석 실험

4.1 실험 환경 및 방법

실제 클라우드 서비스에 대한 리뷰를 대상으로 다양한 기계 학습 알고리즘에 대한 감정 분석 실험을 진행하였다. 실험에 사용된 클라우드 서비스 리뷰 데이터는 Blue Page Dataset[8]으로, SaaS (Software as a Service) 타입의 클라우드 서비스에 대한 약 6,200개의 리뷰로 이루어져 있으며 각 리뷰는 텍스트와 평점 (1~5점 사이의 정수) 으로 구성되어 있다. 본 실험에서는 4~5점짜리 리뷰에는 '긍정적', 1~2점짜리 리뷰에는 '부정적' 레이블을 붙였으며 3점짜리 리뷰는 학습에 사용하지 않았다.

감정 분석 기법으로는 기계 학습 기반의 방법을 이용하였으며, 텍스트 전처리와 기계 학습 알고리즘을 제공하는 Python 라이브러리인 NLTK와 Graphlab Create를 이용하였다. 이를 이용하여 Dummy Text, Stopword 제거 등의 텍스트 전처리 과정을 수행하였다.

기계 학습 알고리즘으로는 지도 학습 (Supervised Learning) 기반 분류 알고리즘인 Decision Tree, Random Forest, Boosted Tree, Logistic Regression, Support Vector Machine을 사용하였다. 또한 학습에 쓰일 Feature를 결정하기 위한 메트릭으로 Term Occurrence, Binary Term Occurrence, TF-IDF를 이용하였다. Term Occurrence (TO, 또는 TF: Term Frequency) 는 각 단어 별 출현 횟수를 말하며, Binary Term Occurrence (BTO) 는 단어의 출현 유무를 말한다. TF-IDF는 TF (TO) 와 IDF의 곱이며 IDF는 D를 문서 집합, t를 단어라고 했을 때 다음 (수식 1)과 같다.

$$IDF(t,D) = \log \frac{|D|}{d \in D : t \in d} \quad (\text{수식 1})$$

위의 세 가지 메트릭 (TO, BTO, TF-IDF) 을 Unigram, Bigram에 대하여 각각 계산하여 Feature로 사용하였다. Unigram은 단어 하나를 단위로 출현 횟수를 세는 방법, Bigram은 단어 두 개를 단위로 세는 방법을 말한다.

성능 평가를 위하여 5-Fold Cross Validation으로 테스트를 진행하였으며 평가 척도로 분류 알고리즘의 평가에 많이 사용되는 F1 Score를 이용하였다. F1 Score를 계산하는 식은 아래 (수식 2)와 같다. True Positive는 실제로 Positive이면서(Actual Positive) 동시에 올바르게 Positive로 예측된(Predicted Positive) 항목을 말한다.

$$F1\ Score = 2 * \frac{Precision * Recall}{Precision + Recall} \quad (\text{수식 2})$$

$$Precision = \frac{True\ Positive}{Predicted\ Positive}$$

$$Recall = \frac{True\ Positive}{Actual\ Positive}$$

4.2 실험 결과

감정 분석 실험을 진행한 결과는 아래 [표 1]과 같다. Boosted Tree 알고리즘에 Bigram 단위로 계산한 Term Occurrence를 Feature로 사용했을 때 가장 좋은 결과를 보였으며 그 외에도 전반적으로 Decision Tree 계열의 앙상블 알고리즘인 Random Forest, Boosted Tree가 뛰어난 성능을 보여주었다.

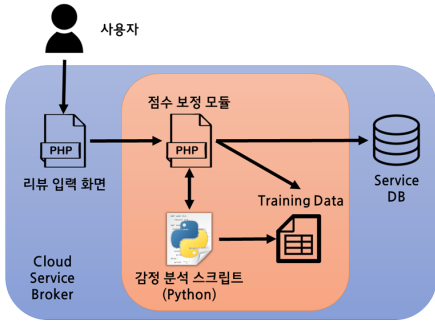
일반적으로 텍스트 분류에서 좋은 성능을 보여주는 것으로 알려진 SVM의 경우 본 실험에서는 다소 좋지 않은 결과를 보였다. 단 본 실험에서 사용한 데이터의 양이 충분히 많다고 보기 어렵고, 특히 '부정적' 으로 분류된 Training Data의 양이 적었다는 점은 감안할 필요가 있다.

[표 1] 감정 분석 실험 결과 (T-Score)

| Feature Algorithm | Unigram+ TO | Unigram+ BTO | Unigram+ TF-IDF | Bigram+ TO | Bigram+ BTO | Bigram+ TF-IDF |
|------------------------|-------------|--------------|-----------------|---------------|-------------|----------------|
| Decision Tree | 0.9818 | 0.9814 | 0.98 | 0.9812 | 0.9808 | 0.9812 |
| Random Forest | 0.9832 | 0.9826 | 0.9818 | 0.9832 | 0.9822 | 0.9828 |
| Boosted Tree | 0.9834 | 0.9834 | 0.9826 | 0.9844 | 0.9836 | 0.983 |
| Logistic Regression | 0.9784 | 0.9792 | 0.9796 | 0.9782 | 0.9778 | 0.9782 |
| Support Vector Machine | 0.9776 | 0.9804 | 0.9766 | 0.978 | 0.978 | 0.9758 |

5. 감정 분석 기반의 사용자 피드백 평가를 적용한 CSB 프로토타입 구현

3장에서 설명한 감정 분석 기반의 사용자 피드백 평가 방법을 적용한 CSB 프로토타입을 구현하였다. CSB에서 사용자 피드백 평가가 동작하는 구조는 아래 [그림 3]과 같다.



[그림 3] 사용자 피드백 평가 동작 구조

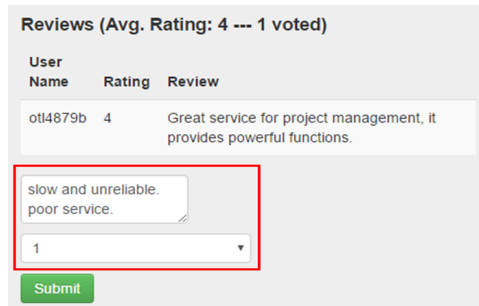
사용자가 서비스 리뷰 입력 화면에서 리뷰 텍스트와 평점을 입력하면 해당 데이터가 점수 보정 모듈에 전달되며, 점수 보정 모듈은 감정 분석을 수행할 Python 스크립트를 호출한다. 감정 분석 스크립트는 일정 주기로 Training Data를 이용해 감정 분석을 위한 기계 학습 기반 분류 모델을 생성 및 갱신하며, 점수 보정 모듈로부터 호출을 받으면 분류 모델을 이용하여 리뷰 텍스트가 긍정적인지, 부정적인지 판단한다. 분류가 완료되면 점수 보정 모듈로 결과를 돌려준다. 점수 보정 모듈은 감정 분석 결과와 사용자가 입력한 평점에 따라 최종적으로 서비스 평가 점수에 반영할 사용자 피드백 점수를 결정하게 된다. 또한 감정 분석이 완료된 리뷰 텍스트는 Training Data Set에 추가하여 학습에 필요한 데이터를 축적한다.

아래에서는 사용자의 피드백에 따라 CSB의 사용자 피드백 점수, 서비스 점수가 변동되는 시나리오를 보인다. 시나리오는 다음과 같이 구성하였다.

- 특정 클라우드 서비스에 대하여 사용자 피드백이 입력되기 전의 서비스 점수를 확인한다.
- 서비스 사용자가 새로운 리뷰를 입력한다.
- 새롭게 등록된 리뷰 텍스트가 감정 분석을 통해 긍정 또는 부정으로 분류되는 것을 확인한다.
- 분류 결과와 평점에 따라 서비스 점수가 변동된 것을 확인한다.

CSB는 서비스를 추천받으려 하는 사용자가 요구사항을 입력하면 그에 따른 서비스의 랭킹과 점수를 보여준다. [그림 4]는 사용자 피드백이 입력되기 전 서비스 점수를 보여주는 화면이다.

사용자가 'Teampage - BASIC' 이라는 클라우드 서비스를 선택하여 이용하고 있다고 가정하자. [그림 5]는 해당 서비스에 대한 리뷰를 남기는 화면이다. 해당 서비스에 대한 기존 리뷰의 개수와 평균 평점, 각 리뷰의 내용과 평점이 표시되며 아래에는 사용자가 리뷰 텍스트와 평점을 입력할 수 있는 인터페이스가 존재한다.



[그림 12] 리뷰 입력 화면

| Recommended Service List | | |
|--------------------------|-------------------------------|-----------------|
| Service ID | Service Name | Service Score |
| 9 | Twproject - enterprise server | 364.64009834768 |
| 6 | Teampage - BASIC | 353.73664631188 |
| 5 | 10000ft - Up to 100 | 347.54461574405 |
| 8 | TargetProcess - Standard | 338.4105245218 |
| 1 | Easy Project - PRO | 334.89728756408 |
| 3 | Assembla - Professional 1 | 324.92415803065 |
| 7 | InLoox Now! - Web User 10 | 323.24471342444 |
| 4 | AceProject - Standard | 295.91545944845 |
| 2 | Zoho Projects - FREE | 290.79674831629 |
| 10 | Smartsheet - Basic | 238.78739585452 |

[그림 4] 사용자 피드백 입력 전 서비스 점수 표시 화면

‘Submit’ 버튼을 클릭하여 리뷰를 남기면 CSB는 해당 리뷰에 대한 감정 분석 결과를 데이터베이스에 저장한다. [그림 6]은 각 리뷰별 감정 분석 결과와 사용자 피드백 점수를 표시하고 있는 화면이다. 기존에 존재했던 리뷰는 Positive로, 새로 입력된 리뷰는 Negative로 분류된 것을 확인할 수 있다. 사용자 피드백 점수는 [그림 2]에서 설명한 방법을 적용하여 계산하였는데, 본 프로토타입에서는 기본적으로 (유저가 입력한 평점 * 10)을 기본 점수로 부여하고 감정 분석 결과에 따라 보정 점수를 추가로 부여하였다. 예를 들어 기존에 존재했던 평점 4점짜리 리뷰의 경우 기본 점수 40점에 Positive로 분류되었으므로 보정 점수 +5점을 받아 사용자 피드백 점수는 45점이 된다. 평점 1점짜리 리뷰의 경우 Negative로 분류되었지만 1점은 대부분 Negative일 것이라 일반적으로 생각할 수 있으므로 Negative일 경우의 보정 점수는 부여하지 않았다.

부정적인 리뷰가 새로 등록되었으므로 평균 사용자 피드백 점수가 감소하였고 이는 서비스 점수에 악영향을 미치게 된다. [그림 7]은 리뷰가 입력된 후 새로운 사용자가 앞의 [그림 4]와 같은 요구사항을 입력했을 때의 서비스 추천 결과를 보여준다. ‘Teampage - BASIC’ 서비스의 점수가 20점 감소하였고, 서비스 점수가 감소함에 따라 랭킹도 3단계 하락한 것을 확인할 수 있다.

6. 결론 및 향후 연구

본 논문에서는 CSB에서 클라우드 서비스를 평가할 수 있는 방법 중 하나로 감정 분석 기반의 사용자 피드백을 이용한 서비스 평가 방법을 제안하였다. 기존의 평점 기반 사용자 피드백의 단점을 보완하기 위해 감정 분석 결과를 이용한 보정 방법을 제시하였으며, 다양한 기계 학습 알고리즘을 적용하여 실제 클라우드 서비스 리뷰에 대한 감정 분석을 수행한 실험 결과를 제시하였다. 또한 제안한 서비스 평가 방법을 적용한 CSB 프로토타입을 구현하였으며, 사용 시나리오를 통해 실제 사용자 피드백을 이용한 서비스 평가의 사례를 제시하였다.

향후 연구로는 클라우드 서비스를 다각도에서 평가 및 추천하는 프로세스를 정립하고, 해당 프로세스를 적용하여 CSB 프로토타입을 발전시킬 계획이다.

| service_id | user_id | rating | sentiment | score | context |
|------------|----------|--------|-----------|-------|---|
| 6 | otl4879b | 4 | positive | 45 | Great service for project management, it provides |
| 6 | newbie1 | 1 | negative | 10 | slow and unreliable. poor service. |

[그림 13] 감정 분석 결과 및 사용자 피드백 점수 확인

| Service ID | Service Name | Service Score |
|------------|--------------------------------|-----------------|
| 9 | Twoproject - enterprise server | 364.64009834768 |
| 5 | 10000ft - Up to 100 | 347.54461574405 |
| 8 | TargetProcess - Standard | 338.4105245218 |
| 1 | Easy Project - PRO | 334.89728756408 |
| 6 | Teampage - BASIC | 333.73664631188 |
| 3 | Assembla - Professional 1 | 324.92415803065 |
| 7 | InLoox Now! - Web User 10 | 323.24471342444 |
| 4 | AceProject - Standard | 295.91545944845 |
| 2 | Zoho Projects - FREE | 290.79674831629 |
| 10 | Smartsheet - Basic | 238.78739585452 |

[그림 14] 사용자 피드백 입력 후 서비스 점수 표시 화면

참 고 문 헌

- [1] NIST Cloud Computing Standards Roadmap. https://www.nist.gov/sites/default/files/documents/itl/cloud/NIST_SP-500-291_Version-2_2013_June18_FINAL.pdf
- [2] Ja-Ryoung Choi, Jungsoo Hwang and Soon-Bum Lim, "Design of User Feedback Interface for Dynamic Updating of E-Book Content", International Journal of Multimedia and Ubiquitous Engineering, Vol. 10, No. 11, pp. 359-368, 2015.
- [3] Emitza Guzman and Walid Maalej, "How Do Users Like This Feature? A Fine Grained Sentiment Analysis of App Review", IEEE 22nd International Requirements Engineering Conference, pp. 153-162, 2014.
- [4] 이태원, 홍태호, "Support Vector Machine을 이용한 온라인 리뷰의 용어기반 감성분류모형", Information Systems Review, 제 17권 1호, pp. 49-64, 2015.
- [5] Lie Qu, Yan Wang and Mehmet A Orgun, "Cloud Service Selection Based on the Aggregation of User Feedback and Quantitative Performance Assessment", IEEE International Conference on Services Computing, pp. 152-159, 2013.
- [6] <https://youtube.googleblog.com/2009/09/five-stars-dominate-ratings.html>
- [7] Walaa Medhat, Ahmed Hassan and Hoda Korashy, "Sentiment analysis algorithms and applications: A survey", Ain Shams Engineering Journal, Vol. 5, No. 4, pp. 1093-1113, 2014.
- [8] Asma Musabah Alkalbani, Ahmed Mohamed Ghamry, Farookh Khadeer Hussain and Omar Khadeer Hussain, "Blue Pages: Software as a Service Data Set", IEEE International Conference on Broadband and Wireless Computing, Communication and Applications, pp. 269-274, 2015.



윤 동 규

2016년 부산대학교 정보컴퓨터공학부 졸업(학사) 2016년~현재 부산대학교 전기전자컴퓨터공학과 석사과정. 관심분야는 소프트웨어 개발 프로세스, 클라우드 컴퓨팅, 기계 학습 등



김 응 수

2016년 부산대학교 정보컴퓨터공학부 졸업(학사) 2016년~현재 부산대학교 전기전자컴퓨터공학과 석사과정. 관심분야는 클라우드 컴퓨팅, 소프트웨어 아키텍처 등



박 준 석

2002년 부산대학교 컴퓨터공학과 졸업(석사) 2012년 부산대학교 컴퓨터공학과 졸업(박사) 2014년~현재 부산대학교 물류혁신네트워킹연구소 연구교수. 관심분야는 소프트웨어 프로덕트 라인 공학, 소프트웨어 아키텍처 등



염 근 혁

1985년 서울대학교 계산통계학과 졸업(학사), 1992년 Univ. of Florida 컴퓨터공학과 졸업(석사), 1995년 Univ. of Florida 컴퓨터공학과 졸업(박사), 1996년~현재 부산대학교 전기컴퓨터공학부 교수. 관심분야는 소프트웨어 아키텍처, 센서 네트워크 기반 상황인식 미들웨어, 빅 데이터 등