

# 심층 강화학습 알고리즘 기반의 감성 데이터 분석 정확성 향상 모델 설계

정세훈\*, 김종찬\*\*, 조경호\*\*\*, 박준\*\*\*, 김준영\*\*\*, 심춘보\*\*\*

\*영산대학교 연계전공학부

\*\*순천대학교 컴퓨터공학과

\*\*\*순천대학교 정보통신·멀티미디어공학과

e-mail : shjung@ysu.ac.kr, cbsim@sunchon.ac.kr

## A Study of Sentimental Data Analysis Accuracy Improvement Model based on Deep RL Algorithm

Se-Hoon Jung\*, Jong-Chan Kim\*\*, Kyeong-ho Jo\*\*\*, Joon Park\*\*\*,

Joon-young Kim\*\*\*, Chun-Bo Sim\*\*\*

\*School of Major Connection, YoungSan University

\*\*Dept of Computer Engineering, Sunchon University

\*\*\*School of Information Communication and Multimedia, Sunchon University

### 1. 연구 필요성 및 문제점

4차 산업혁명이 도래하고 빅데이터 시대에서 소프트웨어의 중요도는 매우 중요하다. 특히 4차 산업혁명 시대에 모바일과 모바일에서 생성되는 데이터는 모든 산업의 주요 근간이 되고 있다. 모바일에서 생성되는 데이터인 SNS 데이터는 반정형, 비정형 데이터로 인간의 감성을 분석할 수 있는 중요한 자료가 된다. 현재 감성분석에 대한 연구가 지속적으로 증가하고 있는 실정이다. 기존 감성 분석은 긍정과 부정에 대한 결과를 출력하는 양극성을 가지는 분석 성질을 포함하고 있다. 감성분석은 기계학습을 기반으로 분석을 처리한다. 레이블 설정 후 분석을 진행하는 지도학습과 비레이블링의 분류와 분석을 진행하는 비지도학습이 대표적인 감성분석 알고리즘이다. 그러나 현재 다양한 연구에서 기계학습 알고리즘 처리 단계중 전처리(Preprocessor) 과정을 통해 분석 비용을 감소시키는 연구가 진행되고 있지만, 숫자 변환을 위한 비정형 데이터의 전처리 과정은 불필요한 비용이 증가되고 문제점이 발생되고 있다. 이에 본 연구에서는 비레이블링된 SNS 데이터를 기준으로 클러스터 알고리즘과 스스로 학습할 수 있도록 심층 강화학습[1]기반으로 감성 분석 데이터의 다양한 결과를 판단하는 감성 데이터 분석 모델을 제안한다.

### 2. 연구내용과 방법

그림 1은 본 논문에서 제안하는 심층 강화학습 알고리즘 기반의 감성 데이터 분석 모델의 구조도이다. 본 논문에서 제안하는 감성 데이터 분석 모델의 아키텍처는 기본적으로 K-means 알고리즘의 문제점을 보완하는 새로운 데이터 분류 알고리즘과 DQN에서 사용되는 네트워크 출력인 Q-Value를 포함한다. 데이터 클러스터링 아키텍처는

크게 2개의 레벨로 구분하여 설계한다. 강화된 클러스터 알고리즘을 통해 텍스트 감성 데이터를 분류하는 Clustering Level, 군집화된 감성 데이터를 스스로 학습하기 위한 Reinforcement Learning Level로 구분하여 설계를 진행한다. Reinforcement Learning Level은 Off-policy 방식과 Q-Table을 변형시킨 A-Deep Q-Learning을 적용한 클러스터링의 감성 분석 데이터의 학습 모델을 포함한다.

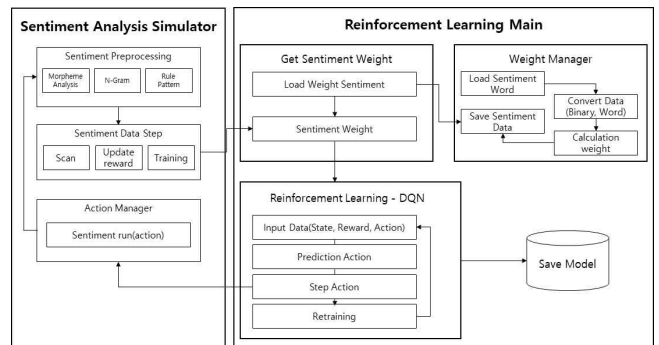


그림 1. 제안하는 감성 분석 모델 구성도

### 3. 결론 및 향후 연구

본 연구에서는 비레이블링된 감성 데이터를 기준으로 기존 DQN에서 사용되는 네트워크 출력인 Q-Value를 활용한 심층 강화학습과 클러스터 분석기반의 감성 데이터 분석을 통해 다양한 감성 데이터 분석 모델의 설계를 제안한다.

### 참고문헌

[1] T. D. Kulkarni, A. Saeedi, S. Gautam, "Deep Successor Reinforcement Learning," arXiv preprint arXiv:1606.02396 (2016), pp. 1-10, 2016.

※ 이 논문은 2019년도 과학기술정보통신부의 재원으로 한국연구재단의 지원을 받아 수행된 연구임 (No.2019R1G1A1002205).