

인터넷 기사의 헤드라인과 본문 사이의 연관성 분석 소프트웨어 모듈: CHIMERA

송성한*, 최정환*, 강민구*, 유철중*

A Software Module That Analyzes the Relationship Between Headline and Content of the Web Article: CHIMERA

Seounghan Song*, Jeongwhan Choi*, Mingu Kang*, and Cheoljung Yoo*

요 약

인터넷 매체를 통한 뉴스의 소비가 급격하게 증가하는 시대에서 언론들은 저널리즘의 역할을 적절히 실천하고 있지 않기 때문에 헤드라인을 통해 뉴스 소비자가 예상한 기사의 맥락과는 다르게 본문이 전개되는 인터넷 기사들로 인한 문제점들이 만연하다. 이 문제를 해결하기 위해서 인터넷 기사가 탈 맥락의 특징을 가진 기사인지 아닌지를 구분할 수 있는 판단 기준을 제공하기 위한 연구가 이루어져 왔으나 정확한 예측 결과를 제공하지 못했다. 이후 인공지능망을 이용한 추후 연구가 진행되었고 기존의 방법들보다 더 좋은 성능을 보였다. 이 논문에서는 인공지능망을 이용한 선행된 연구들을 소개하고 CHIMERA 소프트웨어 모듈에 사용된 인공지능망의 예측 정확도를 측정하였다. 그 결과 측정된 인공지능망의 일부 성능은 서비스로 제공되기에 적합함을 보였다.

Abstract

In the era of rapid growth in the consumption of news through Internet media, the problems caused by Internet news whose text is not in line with the context of news consumer's predictions through headlines are pervasive because the media is not properly practicing the role of journalism. In order to solve this problem, efforts were made at home and abroad to provide criteria for determining whether an Internet news is an news with deviating contextual features, but failed to provide accurate analysis results. Later studies using artificial neural networks were conducted and performed better than conventional methods. In this paper, we introduced the preceding studies using artificial neural networks and measured the predicted accuracy of artificial neural networks used in CHIMERA software modules. As a result, some of the measured performance of the artificial neural network was shown to be inadequate for service.

Key words

fake news, artificial neural network, stance analysis, headline effect, clickbait, de-contextual, multi-layer perceptron, tf-idf

* 전북대학교 소프트웨어공학과

1. 서 론

기존의 뉴스 소비 시장은 언론에게 진실성 및 진정성을 기준으로 뉴스를 작성 및 배포하는 저널리즘의 역할을 부여한다. 이로써 언론은 자신이 작성한 뉴스를 배포하기 전에 언론 내에 갖추어진 검토 체계를 거쳐 배포 여부를 결정하게 된다. 하지만 이 검토 체계는 언론과 독립적이지 못하다. 또한, 검토 체계에서 이루어지는 검토 과정이 뉴스 소비자에게 표면적으로 드러나지 않는다. 동시에 인터넷의 발달로 인해 인터넷상에 존재하는 언론의 수가 급격하게 증가하고 있다. 새로이 등장한 인터넷 언론은 수익을 창출하기 위해 뉴스 소비자들을 겨냥한 뉴스를 주로 생산한다[1]. 하지만 뉴스 검토 체계에 대해서 뉴스 소비자들이 쉽게 확인할 수 없다는 심각한 문제점이 있다.

이로 인해 헤드라인과 본문과의 연관성이 떨어지고 자극적인 헤드라인으로 인터넷 뉴스 소비자의 클릭을 유도하는 인터넷 뉴스 기사인 페이크 뉴스(Fake News)에 대해서 제대로 된 검토와 규제가 이루어지고 있다고 판단하기 어렵다.

언론사의 관점에서 TV 매체를 통한 기사의 전달은 제한된 방송 송출 시간 내에서 중점적으로 다루어질 기사를 압축하여 설명하기 때문에 뉴스의 헤드라인이 본문을 대변하도록 한다. 하지만 인터넷 뉴스의 경우에는 기사의 헤드라인이 본문을 대변할 필요는 없다. 인터넷 상에서의 뉴스 소비는 전적으로 뉴스 소비자의 선택에 달려있고 시간도 제한되어 있지 않기 때문이다.

이러한 차이점을 구분하지 못한 채 인터넷 웹페이지에 게시된 기사의 헤드라인만을 읽고 마치 확인된 사실로 여기는 사람들이 늘어나고 해당 기사를 쉽게 배포하는 등의 문제점이 발생 되어 사회의 혼란을 야기하고있다[2]. 또한 표 1에 제시된 것과 같이 소비자의 뉴스 소비 경로가 종이 신문으로부터 모바일 또는 인터넷 환경으로 이동하고 있다는 사실을 알 수 있으며, 이로 인해 인터넷 환경에서 페이크 뉴스가 현재 또는 미래에 미칠 영향은 더욱 클 것이라 예상된다.

본 논문에서는 위 문제점 해결을 위한 방법으로 인공지능경망을 소개한다. 현대의 인공지능경망 연구는

사람이 쉽게 할 수 없는 일을 기계에 일임하여 해결하고자 하는 목표를 가지고 있다. 많은 양의 기사들 속에서, 사람이 기사의 헤드라인과 본문이 가지는 맥락을 비교하고 스탠스(Stance)를 예측하는 방법은 시간적으로 한계가 있다. 하지만, 주어진 페이크 뉴스 데이터를 기반으로 학습된 인공지능경망을 이용하여 스탠스를 예측한다면 기존의 뉴스 검토 체계가 가지는 시간적 한계점을 극복할 수 있다.

표 1. 국내 뉴스 소비 경로 추이

Table 1. Trends in Domestic News Consumption

년(연도)	모바일(%)	인터넷(%)	종이신문(%)
2012	47.4	62.7	40.9
2013	55.3	65.4	33.8
2014	59.6	67.1	30.7
2015	65.4	69.7	25.4
2016	70.9	73.7	20.9

II. 관련 연구

본 논문에서는 위의 문제점을 인공지능경망을 통해 해결하려는 시도들을 소개하고 CHIMERA 소프트웨어 모듈에 사용된 기반 인공지능경망의 페이크 뉴스 예측 정확도를 검토한다. 인공지능경망을 도입하여 시행된 국내 및 국외의 사례는 표 2에 정리하였다.

표 2. 국내 및 국외 페이크 뉴스 문제 해결 시도

Table 2. Attempts to Resolve Fake News (domestic, oversea)

국내	국외
ESTSoft	SOLAT in the SWEN (Cisco)
IWAZ	Athene (UKP Lab)
Nuaa	UCL Machine Reading

국내의 경우 페이크 뉴스를 판별하는 인공지능 R&D 챌린지에서 ESTSoft사가 개발에 참여하여 만들어진 인공지능경망이 1등 수상을 하였다. 하지만 사용자를 위한 서비스까지는 도달하지 못한 것으로 조사되었다[3].

국외의 경우 동일한 문제를 해결하기 위하여 FNC(Fake News Challenge) 대회를 개최하여, 인공지능경망 공모전을 진행한 사례가 있다[4]. 해당 공모전에서는 미리 페이크 뉴스에 대한 데이터가 제공되었고 동일한 테스트-데이터셋을 공모전에 참여한

인공신경망에 입력하여 얻은 결과로 성능을 비교하였다. 성능 비교표는 아래의 표 3과 같다.

표 3. 국외 인공신경망 성능 비교표
Table 3. Performance Comparison of ANN

프로젝트 명	예측 정확도(%)
SOLAT in the SWEN (Cisco)	82.02
Athene (UKP Lab)	81.97
UCLMR-fnc (UCL Machine Reading)	81.72

성능비교표에 따르면 Cisco사에서 진행한 프로젝트인 SOLAT in the SWEN[5]이 82.02%로 가장 높은 예측 정확도를 보였다. 하지만 제시된 인공신경망의 복잡도 및 향후 코드의 개선을 통한 페이크 뉴스 예측 정확도의 향상 가능성, 코드의 유지 보수성을 포함하여 고려하였을 때 최종적으로 UCL Machine Reading에서 제안한 프로젝트[6]의 인공신경망을 CHIMERA 소프트웨어 모듈의 기반 인공신경망으로 선정하였다.

III. 인터넷 기사 분석을 위한 소프트웨어 모듈

CHIMERA 소프트웨어 모듈 아키텍처는 그림 1과 같다. 서버는 유저 인터페이스를 통해 사용자의 현재 URL을 받아온다. 주어진 URL로부터 스탠스 분석 모델에 입력 가능한 형식으로 데이터를 가공한다. 이를 입력하여 분석 모델이 예측한 결과를 유저 인터페이스를 통해 사용자에게 제공한다.

CHIMERA 소프트웨어 모듈의 스탠스 분석 모델

은 그림 2와 같다. 이는 하나의 은닉층으로 이루어진 다층 퍼셉트론(Multi Layer Perceptron) 구조이다. FNC에서 제공하는 기사의 헤드라인과 본문 데이터를 이용하여 각각의 TF(Term-Frequency) 벡터를 추출하여 특징 벡터(feature-vector)를 만든다. 그리고 헤드라인과 본문의 TF-IDF 벡터 간의 코사인 유사도 벡터를 피쳐 벡터에 추가한다. 이를 토대로 ReLU(Rectified Linear Unit) 활성화 함수가 적용된 은닉층에 입력된다. 마지막으로 출력 가능한 4가지 스탠스인 'agree'(동의), 'disagree'(비동의), 'discuss'(토의), 'unrelated'(관련없음) 중 가장 높은 확률을 가진 스탠스로 예측을 한다.

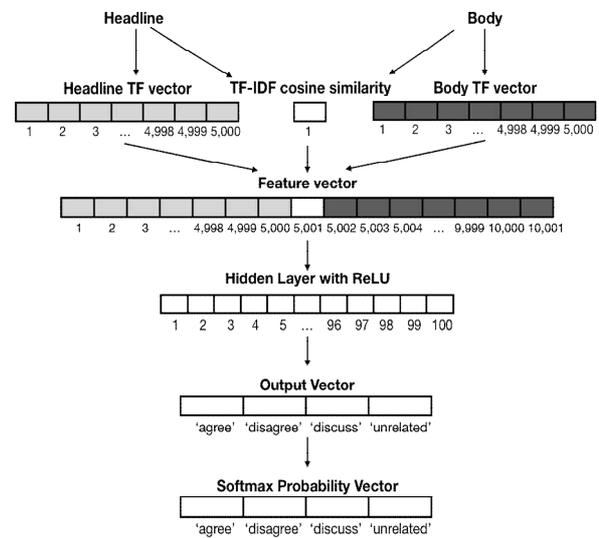


그림 2. 스탠스 분석 모델
Fig. 2. Stance Analysis Model

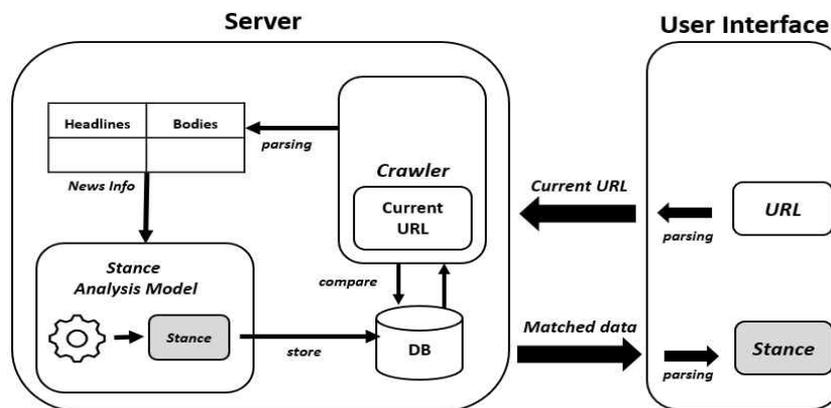


그림 1. CHIMERA 모듈 아키텍처
Fig. 1. CHIMERA Module Architecture

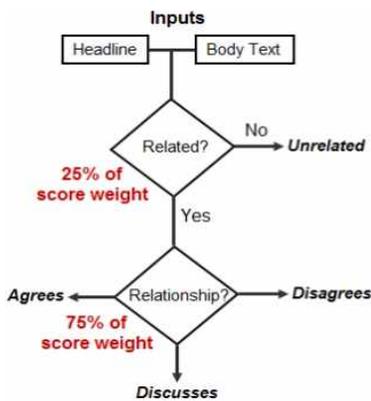


그림 3. 스탠스 탐지 과정
Fig. 3. Stance Detection Process

헤드라인과 본문 사이의 관련성이 없는 경우인 'unrelated'를 제외하고 헤드라인과 본문이 의미하는 바가 일치하면 'agree'를 반환하고, 헤드라인과 본문이 의미하는 바가 일치하지 않는다면 'disagree'를 반환하게 된다. 또, 헤드라인이 의미하는 바에 대해서 본문이 다루고 있지만 뚜렷한 주장이 없다면 'discuss'를 반환한다. 그림 3은 이러한 과정을 보여주고 있다.

IV. 토 의

FNC에서 제공하는 테스트-데이터 세트에 대한 UCLMR-fnc 프로젝트의 인공지능망의 성능은 아래의 표 4와 같다.

표 4. UCLMR-fnc 인공지능망 성능
Table 4. Performance Measurement of UCLMR-fnc

True \ Pred.	'agree'	'disagree'	'discuss'	'unrelated'	Overall	% Accuracy
	'agree'	838	12	939		
'disagree'	179	46	356	116	697	6.60
'discuss'	523	46	3,633	262	4,464	81.38
'unrelated'	53	3	330	17,963	18,349	97.90
Overall	1,593	107	5,258	18,455	25,413	88.46

각 스탠스에 대한 예측 정확도가 50% 미만인 스탠스의 데이터를 분석한 결과는 다음과 같다. 'agree' 스탠스에 대한 예측 정확도는 44.04%이고 'agree'로 예측되어야 할 기사가 'discuss'로 다수 측정되었다. 'disagree' 스탠스에 대한 예측 정확도는 6.60%이다. 이는 예측되어야 할 스탠스인 'disagree'가 아닌 'unrelated' 또는 'agree', 'discuss' 등으로 예측이 되었

다. 이를 통해서 CHIMERA 소프트웨어 모듈의 기반으로 사용된 인공지능망의 성능이 향상되어야 한다는 것을 알 수 있다.

V. 결 론

본 논문에서는 인터넷 기사의 헤드라인과 본문 사이의 연관성을 분석하기 위한 CHIMERA 소프트웨어 모듈을 구성하기 위해 사용된 인공지능망의 페이크 뉴스 예측 정확도를 측정하였다.

그 결과, 'agree' 및 'disagree' 스탠스는 각 44.04%, 6.60%의 예측 정확도를 가지고 있어 서비스로 제공하기에 부적합하다고 판단된다. 하지만 'discuss' 및 'unrelated' 스탠스는 각 81.38%, 97.90%의 예측 정확도를 가지고 있어 서비스로 제공하기에 적합하다고 판단된다. 따라서 향후 연구를 통해 각 스탠스에 대한 예측 정확도 향상을 위한 방법을 도출 해보는 것이 필요하다.

참 고 문 헌

- [1] 이선민, “헤드라인 선정성과 기만성이 뉴스 품질과 정서를 매개로 뉴스 이용의도에 미치는 효과”, 2019. 연세대학교 커뮤니케이션대학원
- [2] “SBS ‘세월호 오보’ 부실취재·부적절 데스크링 문제”, <한겨레>, 2017년 5월 15일, <http://www.hani.co.kr/arti/society/media/794841.html> (2019. 05. 20 - 접속날짜)
- [3] “가짜뉴스 꼭 잡아내는 ‘AI 보안관’ 나온다”, <동아일보>, 2017년 12월 11일, [http://www.fakenewschallenge.org/](http://news.donga.com/3/all/20171211/87668182/1(2019. 05. 20 - 접속날짜)
[4] “Fake News Challenge”, <FNC.org>, <a href=) (2019. 03. 22 - 접속날짜)
- [5] Sean Baird, Doug Sibley, Yuxi Pan, June 2017, <https://blog.talosintelligence.com/2017/06/talos-fake-news-challenge.html>
- [6] Benjamin Riedel, Isabelle Augenstein, Georgios P. Spithourakis and Sebastian Riedel, "A simple but tough-to-beat baseline for the Fake News Challenge stance detection task", Cornell University, Computer Science, Computation and Language, May 2018.