**최종 보고서**

****

**T-SA:** **트위터 단어 검색 기반의 트윗 분석**

(19대 대통령 선거)



학 기: 2019-1 학기

과 목: 산학캡스톤디자인

담당교수: 정 현 숙 교수님

팀 명: VI

팀 장: 이석준(20165072)

팀 원: 이윤혁(20165062)

배인규(20165073)

서재익(20144773)

**- 차 례 -**

**제1장. 서론**

**제1절. 개발 동기 및 필요성**

**제2절. 기존 연구**

**제2장. 본론**

**제3장. 결론**

**제4장. 부록**

**제1장. 서론**

**제1절. 개발 동기 및 필요성**

트위터는 트윗에 자신의 의견(생각)을 표출할 수 있는 SNS(Social Network Service)이다. 지금은 페이스북, 인스타그램보다 사용률이 많이 낮아졌지만, 트윗 작성의 편리함과 정보의 빠른 확산성 때문에 사회적 영향은 무시할 수 없다. 계정 생성 시 메일 주소만을 필요로하고, 형식상 휴대폰 인증과정을 거치지만 다른 사용자에게는 비공개되기 때문에 익명성은 다른 SNS에 비해 강한 편에 속한다. 따라서 자신의 의견(생각)을 자신 있게 표현하는 경우가 많아서 사회현상에 대한 다양한 의견(생각)이 표출된 트윗이 많이 생성된다. 트위터의 게시문인 트윗은 한글이든 영문이든, 공백과 기호를 포함해 한 번에 140글자를 작성할 수 있다. 이러한 특징 때문에 다른 사용자의 의견(생각)이 표현된 트윗을 보는 데 필요한 시간이 적게 소요되며, 실시간 이슈가 되는 주제에 대한 내용을 트위터 검색을 통해 다른 SNS(페이스북, 인스타그램)에 비해 빠르고 다양한 결과를 찾아볼 수 있다.

2008년과 2012년 미국 대통령 선거에서 오바마 캠프가 전 세계적인 주목을 받았었다. 오바마 캠프에서는 출마 선언부터 SNS(Social Networking Sites OR Social Network Services)에 공개하였으며, 주요 지지층을 결집하고 정치적 우군으로 동원하는 과정에도 SNS를 활용하였다. 그래서 2008년과 2012년 대통령 선거에서 버락 오바마가 뛰어난 소셜 미디어 운영 방식을 선보이고, 활용하여 긍정적인 결과를 보였기 때문에 버락 오바마는 소셜 미디어 대통령이라는 별명을 가지게 되었다. 또한 버락 오바마 대통령은 선거 캠페인 도중에 '소셜 미디어는 후보들이 저널리즘의 필터 없이 이야기 할 수 있는 플랫폼'이라고 언급도 했다. 그 이후, 2016년 미국 대통령 선거에는 힐러리 클린턴(민주당)과 도널드 트럼프(공화당)가 출마를 하였다. 이때 도날드 트럼프는 1130만 명의 트위터 팔로워와 1056만 명의 페이스북 팔로워를 보유하고 있었고 이는 1000만 명의 트위터 팔로워를 보유하고 있던 전 대통령인 버락 오바마를 훨씬 뛰어넘는 수치였다. 이러한 많은 수의 SNS 팔로워를 통해 도날드 트럼프는 공화당 후보 경쟁 기간 내내 각종 정치 현안과 관련된 메시지를 SNS에 쏟아내며 화제를 몰고 다녔다. 특히 자신의 공약 홍보나 치적을 보여주기보단 상대 후보(힐러리 클린턴)를 노골적으로 비난하고 불법 이민, 인종차별 등의 민감한 문제에 대해 극단적인 발언을 서슴지 않고 내뱉으며 미국 정치권을 충격에 빠트리기도 했었다. 하지만 SNS 속 지지자들은 이러한 도날드 트럼프의 극단적인 발언에 열광했고 그 결과 2016년 대통령 선거에서 당선되었다.

우리나라는 2002년에 대통령 선거에서 인터넷을 적극적으로 활용한 경험을 통해, 2011년 서울 시장 보궐 선거에서 본격적으로 소셜 미디어가 어떻게 선거에 활용될 수 있는가를 보여주었다. 특히 각 후보의 네트워크와 메시지의 전파, 소셜 미디어에서의 투표 참여 독려 등 여러 측면에서 소셜 미디어의 역활에 대한 깊이 있는 분석이 등장하였다.

앞에 나온 미국 대통령 선거(2008년, 2012년, 2016년)와 우리나라의 서울 시장 보궐 선거(2011년)에서 SNS를 이용한 사례를 통해 SNS와 우리나라의 대통령 선거와의 관계성이 있을 수 있다고 예상된다. 그래서 우리는 이번 프로젝트에서 SNS 중 위에 작성한 것처럼 짧은 글로 작성되는 트위터와 우리나라의 대통령 선거 중 19대 대통령 선거를 선택하여 이둘의 관계성을 분석할 예정이다.

**제2절. 기존 연구**

와이즈넛에서 제공하는 트윗 분석 사이트인 트윗 트렌드(http://tweetrend.com)는 약 400만명의 트위터 계정에 대해 과거 2011년 1월 이후의 데이터부터 하루 1천만 트윗을 수집하여 국내 최대의 검색 서비스를 제공하고 있다. 대표적인 서비스로는 리트윗, URL 파싱, 파급력 등 다각적인 분석 결과를 제공한다. 대표적인 예로는 마케팅 등 다양한 분야에서 활용할 수 있도록 원문을 내려받아 볼 수 있는 서비스가 있다. 이를 이용할 수 있는 방법으로는 3가지(비로그인, 무료회원, 유료회원)로 나누어진다. 각각의 방법의 사용 가능한 검색 키워드의 갯수와 검색 기간은 표1과 같다.

|  |  |  |  |
| --- | --- | --- | --- |
|  | 비로그인 | 무료회원 | 유료회원 |
| 검색 키워드 갯수 | 1 | 3 | 3 |
| 검색 기간 | 7일 | 30일 | 무제한 |

**<표 1>** 트윗 트렌드 이용 방법별 검색 권한

콘스탄틴 코브세닌(Konstantin Kovshenin)이 Twitter API를 이용하여 만든 사이트인 팔로우 미(http://foller.me)는 공개된 프로필에 관한 주제, 언급, 해시태그 등 정보를 분석하고 보여준다. 대표적인 서비스로는 요청한 사용자의 프로필과 사용자가 작성한 최신 트윗들을 수집하여, 내용을 분석하고 자주 사용되는 단어(해시태그)들을 출력하여 보여준다.



그림1. tweetrend Site(http://tweetrend.com)