

보건·복지 ISSUE & FOCUS

Korea Institute for Health
and Social Affairs

ISSN 2092-7117
제 296호 (2015-28) 발행일 : 2015. 10. 19

KIHASA 한국보건사회연구원
Korea Institute for Health and Social Affairs

소셜 빅데이터를 활용한 우리나라 청소년의 우울 현황과 위험요인 분석¹⁾

청소년의 우울관련 온라인 문서는 개학 시기인 3월부터 증가하여 5월에 감소하며, 다시 여름방학 후 9월부터 증가하여 11월부터 감소 추세를 보이는 것으로 나타남

청소년의 우울관련 부정적인 감정(위험)을 나타내는 온라인 문서는 2012년 25.2%, 2013년 27.9%, 2014년 24.5%로 나타남

온라인 문서에서 '질병'과 '가정폭력'이 동시에 언급되면 청소년의 우울감정은 위험함

청소년은 온라인상에서 자살생각과 같은 우울한 담론을 언급하기 때문에 온라인상에서 우울의 징후가 예측되면 실시간으로 개입할 수 있는 정부차원의 우울관리 모니터링 체계의 구축이 필요함



송태민
정보통계연구실 연구위원

1. 서론

- 우울증은 정신병리학에서 '현대인의 정신적 감기'라고 불릴 정도로 남녀노소에 관계없이 일반인 모두에게 익숙한 용어이나 우울증의 폐해는 심각함
 - 청소년기의 우울증은 만성적이며 재발적인 경과를 나타낼 뿐만 아니라 대인관계 기능 및 학업성취에 악영향을 미침으로써 지속적인 신체성장을 방해하고, 다른 정신장애의 가능성을 증가시키는 등 중요한 사회적 문제로 인식됨²⁾

1) 본 연구는 보건복지부 수탁과제인 '보건의료 빅데이터 활용을 위한 기본계획 수립 연구'의 일환으로 수행된 것임을 밝힘

2) 권호인. 아동청소년기 우울증의 예방 개입에 대한 고찰. 인지행동치료. 12(1) 21-41, 2012

- 우울이 사회문제가 되는 것은 자살과 연관되어 있기 때문으로, 자살한 사람들은 대개 우울증을 앓고 있으며³⁾, WHO에서도 세계질병부담(GBD) 연구를 통해 우울장애를 21세기에 개입해야 할 가장 중요한 목표 중 하나로 지적함⁴⁾

■ 전 세계 우울증 환자는 계속 증가하고 있으며 우울증에 따른 질병 부담도 증가함

- 전 세계 우울증 환자는 3억 5천만 명에 이르고 이 숫자는 계속 증가할 것으로 예측되며(WHO, 2012)⁵⁾, 우울증에 따른 질병 부담은 2004년 3위로 보고되었으나, 2030년에는 1위가 될 것으로 예측됨(WHO, 2008)⁶⁾
- 우리나라의 2009년 우울증 진료인원은 555,528명(남자: 169,180명, 여자: 386,348명)에서 2013년 664,616명(남자: 208,756명, 여자: 455,860명)으로 약 10만 9천명(19.6%)이 증가하였고, 총 진료비는 2009년 약 2,135억 원에서 2013년 약 2,714억 원으로 5년간 약 579억 원(27.1%) 증가하여 연평균 증가율은 6.2%로 나타남(건강보험심사평가원, 2014. 10. 27. 보도자료)
- 2014년 국내 인구 10만 명당 자살률은 27.3명(자살자수: 13,836명)임. OECD 국가 간 자살률 비교 시 OECD 평균 12.0명에 비해, 한국은 28.7명(2013년 기준)으로 가장 높은 수준임('2014년 사망원인 통계', 통계청, 2015. 9. 23. 보도자료)

■ 우리나라 청소년의 우울증상 경험률은 점차 낮아지고 있으나 여전히 높은 수준에 있음

- 우리나라 청소년의 우울증상 경험률은 2011년 32.8%, 2012년 30.5%, 2013년 30.9%, 2014년 26.7%임⁷⁾

2. 소셜 빅데이터의 활용 및 분석의 필요성

- 세계 각국의 정부와 기업은 SNS를 통해 생산되는 소셜 빅데이터의 활용과 분석을 통해 사회적 문제의 해결과 정부의 정책을 효과적으로 추진할 수 있을 것임
 - 최근 SNS 환경의 규모와 영향력이 점차 증대됨에 따라⁸⁾, 개인의 커뮤니케이션 수단으로 SNS는 긴급 상황과 위기대응에 결정적인 요소로 진화함⁹⁾
 - 트위터, 블로그, 카페, 게시판, 뉴스채널 등 온라인 매체에 남긴 정치, 경제, 사회, 문화에 대한 메시지가 그 시대의 감성과 정서를 파악할 수 있는 원천으로 등장함에 따라 소셜 빅데이터 분석을 통한 정책 수요 예측이 가능하게 됨

3) Han SS, Kang SW, Yoo WK, Phee YG. A Study of the Determinants of Suicidal Ideation among the Elderly in Korea, Health and Social Welfare Review 2009; 29(1): 192-212, (Korean)

4) WHO, The global burden of disease: 2004 update, Geneva 2008, Switzerland

5) <http://www.who.int/mediacentre/factsheets/fs369/en/> 2015.6.2.검색

6) WHO(2008). The global burden of disease: 2004 update

7) 교육부, 보건복지부, 질병관리본부(2014). 제10차(2014년) 청소년건강행태온라인조사 통계

8) Jong-Hwan Kong, Ik-Kun Kim, Myung-Mook Han (2014). Propagation Models for Structural Parameters in Online Social Networks, Journal of Internet Computing and Services 15(1) pp. 125-134

9) Ryu, Hyeon Suk (2013). A Study on Risk Perception and Communication via Social Media. The Korea Institute of Public Administration, Research Report 25(3)

- 소셜 빅데이터의 분석은 사용자가 남긴 문서의 의미를 분석하는 것으로 자연어 처리기술인 주제분석(Text Mining)과 감성분석 기술인 오피니언 마이닝(Opinion Mining)을 실시한 후, 네트워크 분석(Network Analysis)과 통계분석(Statistical Analysis)을 실시함
 - 기존에 실시하던 횡단적 조사나 종단적 조사 등을 대상으로 한 연구는 정해진 변인들에 대한 개인과 집단의 관계를 보는 데에는 유용하나 사이버 상에서 언급된 개인별 문서(버즈: buzz)에서 논의된 관련 정보 상호간의 연관관계를 밝히고 원인을 파악하는 데는 한계가 있음¹⁰⁾
 - 소셜 빅데이터의 분석은 훨씬 방대한 양의 데이터를 활용하여 다양한 참여자의 생각과 의견을 확인할 수 있기 때문에 기존의 오프라인 조사와 병행하면 사회적 문제의 예측을 보다 정확하게 할 수 있음
 - 그러나 본 연구에서 소셜 빅데이터에서의 우울 관련 요인(용어)은 온라인 문서 내에서 발생된 빈도로 정의되기 때문에 기존의 조사 등을 통한 이론적 모형에서의 의미와 다를 수 있으므로 후속 연구에서의 검증이 요구됨

3. 소셜 빅데이터를 활용한 우리나라 청소년 우울 현황과 위험요인 분석

가. 분석대상 및 분석방법

- 본 연구에서는 199개의 온라인 뉴스사이트, 9개의 게시판, 1개의 SNS(트위터), 4개의 블로그, 2개의 카페 등 총 215개의 온라인 채널을 통해 수집 가능한 텍스트 기반의 웹문서(버즈)를 소셜 빅데이터로 정의함
 - 우울 토픽(topic)은 모든 관련 문서를 수집하기 위해 ‘우울’을 사용하였으며, 토픽과 같은 의미로 사용되는 토픽 유사어로는 ‘우울증’ 용어를 사용함
 - 소셜 빅데이터의 수집은 2012. 1. 1.~2014. 12. 31.의 기간에 해당 채널에서 요일, 주말, 휴일을 고려하지 않고 매 시간 단위로 이루어졌고¹¹⁾, 수집된 총 3,703,135건의 텍스트(Text) 문서 중 청소년으로 판단되는 161,581(감정을 표현한 문서는 86,957건임)건의 문서를 본 연구의 분석에 포함함
- 우울 위험을 설명하는 가장 효율적인 예측모형을 구축하기 위해 데이터마이닝의 연관규칙과 의사결정나무 분석, 그리고 시각화 분석을 사용함
 - 연관규칙의 분석 알고리즘은 선형적 규칙(a priori principle)¹²⁾을 사용하였고, 의사결정나무 형성을 위한 분석 알고리즘은 CHAID(Chi-squared Automatic Interaction Detection)를 사용함
 - 기술 분석, 다중응답 분석, 의사결정나무 분석에는 SPSS 23.0을 사용하였고, 연관 분석과 시각화에는 R version 3.2.1을 사용함

나. 분석결과

- 우울 관련 감성분석
 - 우울 관련 감성은 감성어 사전을 개발하여 긍정(‘스트레스 받지 않다, 우울증 퇴치하다, 행복 넘치다’

10) 송주영 · 송태민(2014). 소셜 빅데이터를 활용한 북한 관련 위험인식 요인 예측. 국제문제연구, 가을, pp. 209-243

11) 본 연구를 위한 소셜 빅데이터의 수집은 한국어 트위터를 대상(수집기간 중 총 트위터 문서: 3,099,720,203건)으로 하였으며 ‘(주)SK텔레콤 스마트인사이트’에서 수행함

12) 본 연구의 연관분석의 평가는 지지도 0.001, 신뢰도 0.01을 기준으로 측정함

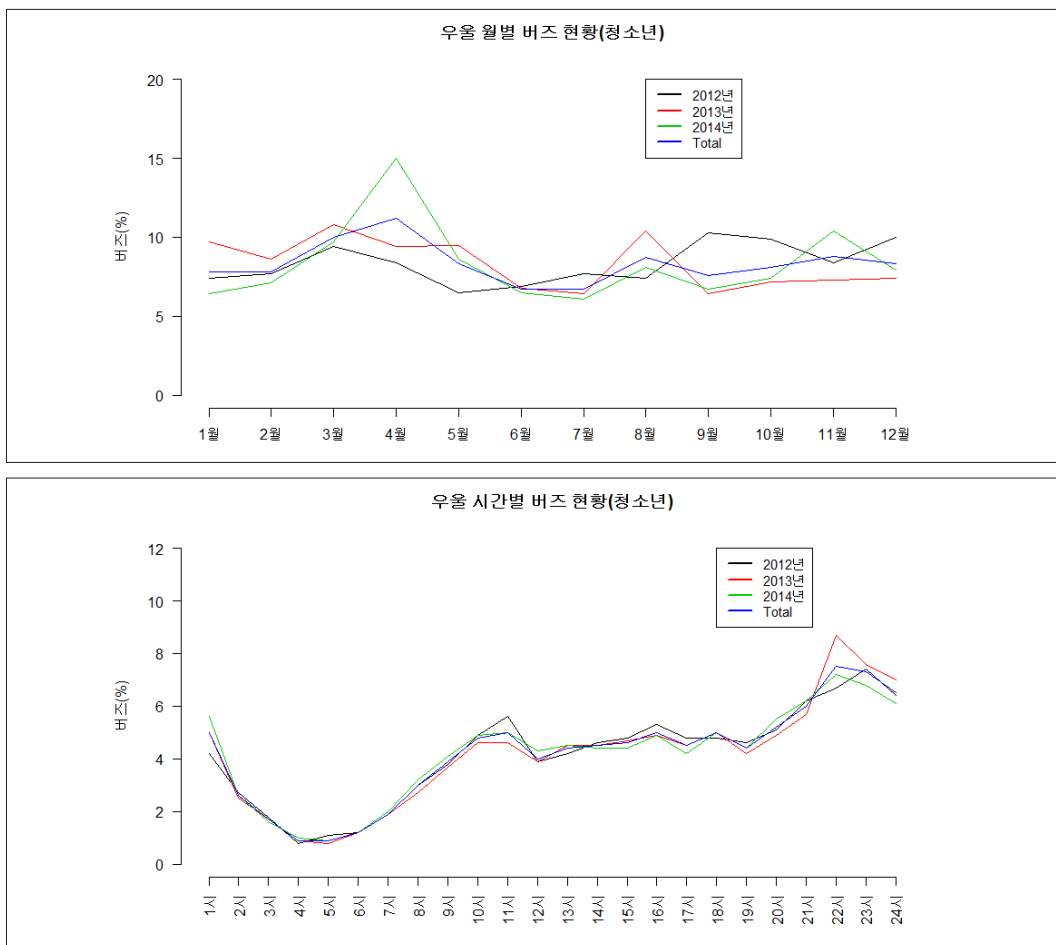
등), 보통(‘한국인 우울하다, 청소년 우울하다’ 등), 부정(‘친구 자살하다. 스트레스 심각하다. 왕따 심각하다’ 등)으로 구분하여 감성분석을 실시함

- 따라서 본고에서 우울의 긍정적 감정은 ‘스트레스 받지 않아 행복하다는 의미’이며, 부정적 감정¹³⁾은 ‘스트레스가 심각하여 우울하다는 의미’를 나타냄
- 우울에 대한 증상은 주제분석¹⁴⁾을 통하여 호흡기 장애, 불안, 공포, 스트레스 등 30개 요인으로 그룹화 하였고, 우울에 대한 원인은 주제분석을 통하여 성격, 질환, 스트레스 등 10개 요인으로 그룹화 하였으며, 우울에 대한 예방은 정신예방, 학교도움 등 6개 요인으로 그룹화 하였음

■ 청소년 우울 관련 문서(버즈) 현황

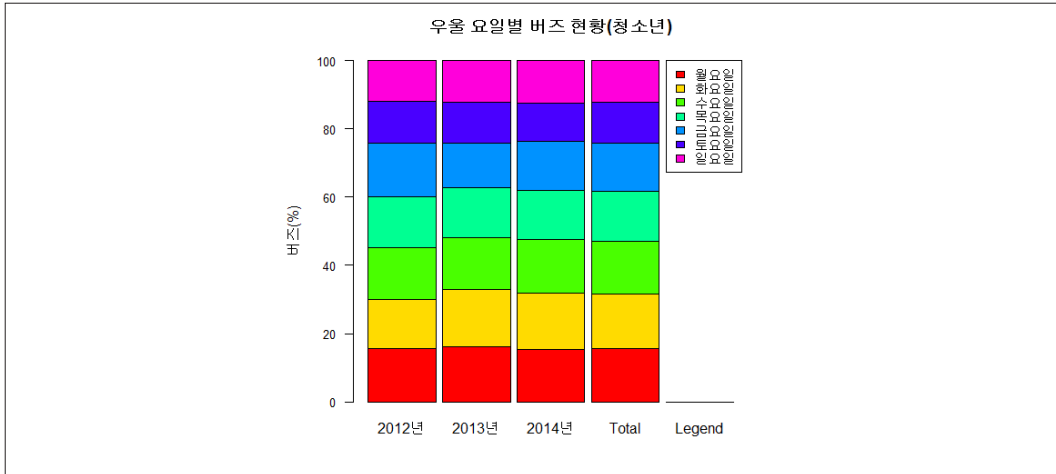
- [그림 1]과 같이 우울과 관련된 문서(버즈)는 개학 시기인 3월부터 증가하여 5월에 감소하며, 여름방학 이후 9월부터 증가하여 11월에 감소 추세를 보이는 것으로 나타남
- 우울과 관련된 문서는 등교 시간인 8시부터 증가하여 11시 이후 감소하며, 다시 20시 이후 증가하여 23시 이후 감소 추세를 보이는 것으로 나타남
- 우울과 관련된 문서는 월요일, 화요일, 수요일에 높은 추이를 보이는 반면, 주말에는 감소하는 것으로 나타남

[그림 1] 청소년 우울 관련 문서(버즈)의 월별, 시간별, 요일별 추이



13) 본 연구의 우울의 부정적 감정은 '청소년 건강행태 온라인조사'에서 청소년 우울증상 경험률의 의미로 사용함

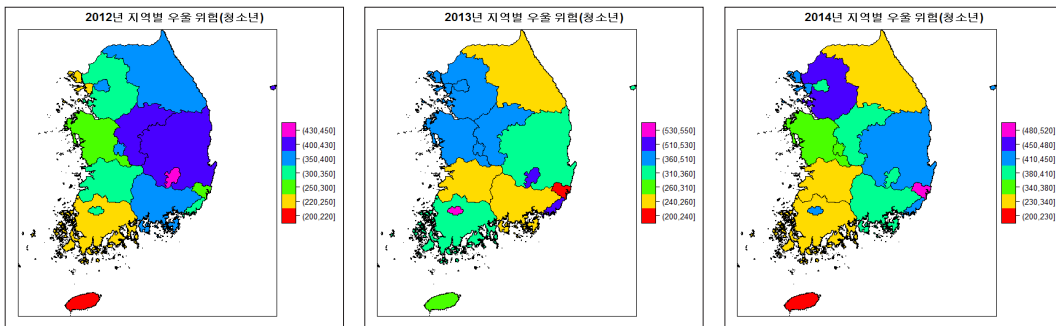
14) 우울의 증상, 원인, 예방 등 주제 분류는 '정혜실(2015). 청소년 우울 관련 소셜미디어 수집과 분석을 위한 온톨로지 개발 및 평가. 서울대학교 대학원 석사학위논문' 발표예정인 연구결과를 적용함



■ 청소년 우울 관련 부정 감정의 지역별 추이

- [그림 2]와 같이 우울에 대한 부정적 감정은 2012년에는 대구광역시와 광주광역시가 가장 높게 나타났고, 제주특별자치도가 가장 낮게 나타났음. 2013년에는 광주광역시가 가장 높고, 울산광역시가 가장 낮으며, 2014년에는 울산광역시가 가장 높고 제주특별자치도가 가장 낮게 나타남

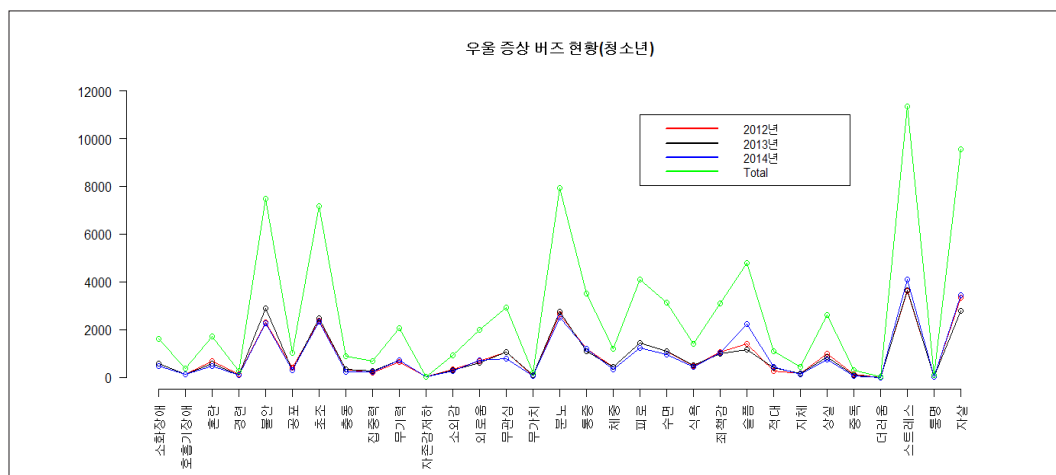
[그림 2] 우울 관련 위험의 지역별 추이



■ 청소년 우울 증상 문서(버즈) 현황

- [그림 3]과 같이 연도별 우울 증상의 변화는 스트레스, 자살생각, 분노, 불안, 초조 등의 순으로 높은 것으로 나타남

[그림 3] 연도별 우울증상 문서(버즈) 현황





2012년 2013년 2014년 Total

- <표 1>과 같이 우울과 관련하여 부정적인 감정(위험)을 나타내는 온라인 문서(버즈)는 25.9%로 우울에 대한 부정적 온라인 문서가 긍정적인 문서보다 높은 것으로 나타남
 - 우울 관련 채널 중 트위터 문서는 61.5%로 나타남
 - 우울관련 원인으로서는 자살생각, 성격, 질환, 성, 스트레스 등의 순으로 나타남
 - 우울관련 예방으로는 가족도움, 사후관리, 태도변화, 신체활동 강화 등의 순으로 나타남

<표 1> 청소년 우울 관련 문서(버즈) 현황

| 구분 | 항목 | N(%) | 구분 | 항목 | N(%) |
|----|--------|--------------|-------|--------------|--------------|
| 감정 | 긍정 | 13,451(15.5) | 연도 | 2012년 | 45,840(28.4) |
| | 보통 | 50,987(58.6) | | 2013년 | 56,017(34.7) |
| | 부정 | 22,519(25.9) | | 2014년 | 59,724(37.0) |
| | 계 | 86,957 | | 계 | 161,581 |
| 채널 | 트위터 | 99,431(61.5) | 원인 | 성격 | 6,453(14.7) |
| | 블로그 | 16,038(9.9) | | 질환 | 5,609(12.7) |
| | 카페 | 23,117(14.3) | | 스트레스 | 4,801(10.9) |
| | 게시판 | 12,509(7.7) | | 성 | 4,881(11.1) |
| | 뉴스 | 10,486(6.5) | | 중독 | 1,708(3.9) |
| | 계 | 161,581 | | 자살생각 | 12,102(27.5) |
| 예방 | 정신예방 | 831(2.0) | | 12,102(27.5) | |
| | 학교도움 | 1,792(4.3) | 주요사건 | 1,674(3.8) | |
| | 가족도움 | 15,980(37.9) | 학교폭력 | 3,877(8.8) | |
| | 사후관리 | 10,493(24.9) | 경제적상황 | 662(1.5) | |
| | 신체활동강화 | 5,423(15.3) | 가정폭력 | 2,239(5.1) | |
| | 태도변화 | 6,598(15.7) | 계 | 44,006 | |
| | 계 | 42,117 | | 44,006 | |

■ 청소년 우울 관련 연도별 감정 변화

- <표 2>와 같이 우울관련 부정적인 감정(위험)을 나타내는 버즈는 2012년 25.2%, 2013년 27.9%, 2014년 24.5%로 부정 감정의 경우 청소년건강행태 온라인조사의 우울감 경험을 추이와 비슷한 것으로 나타남

〈표 2〉 청소년 연도별 우울 감정

(단위: %)

| 연도 | 긍정 | 보통 | 부정 | |
|-------|------|------|-----------------------|---------------------|
| | | | 소셜 빅데이터 ¹⁾ | 교육부 외 ²⁾ |
| 2012년 | 16.8 | 57.9 | 25.2 | 30.5 |
| 2013년 | 15.9 | 56.1 | 27.9 | 30.9 |
| 2014년 | 14.0 | 61.5 | 24.5 | 26.7 |
| 계 | 15.5 | 58.6 | 25.9 | 29.4 |

1) 소셜 빅데이터의 우울에 대한 부정(위험) 감정

2) 교육과학기술부, 보건복지부, 질병관리본부(2014), 제10차 청소년건강행태온라인조사 통계

■ 청소년 우울 원인에 대한 연관성 분석

- 〈표 3〉과 같이 우울 원인에 대한 연관성 예측에서 가장 신뢰도가 높은 연관규칙은 '{질병, 가정폭력}=>{부정}'이며, 세 변인의 연관성은 지지도 0.006, 신뢰도는 0.737, 향상도는 2.845로 나타남. 온라인 문서(버즈)에서 질병과 가정폭력이 언급되면 우울의 감정이 부정적(위험)일 확률이 약 73.7%이며, 질병과 가정폭력이 언급되지 않은 버즈보다 우울에 대해 위험한 확률이 2.85배 높아지는 것으로 나타남

〈표 3〉 청소년 우울 원인 연관규칙

| 규칙 | 지지도 | 신뢰도 | 향상도 |
|------------------------------------|-----------|----------|----------|
| {질병, 가정폭력} => {부정} | 0.006405 | 0.736772 | 2.845043 |
| {성격, 질병, 스트레스, 가정폭력} => {부정} | 0.001195 | 0.732394 | 2.828137 |
| {성격, 질병, 스트레스, 자살생각, 가정폭력} => {부정} | 0.001069 | 0.726562 | 2.805617 |
| {질병, 스트레스, 자살생각, 가정폭력} => {부정} | 0.001402 | 0.726190 | 2.804180 |
| {성격, 질병, 스트레스, 중독, 자살생각} => {부정} | 0.001207 | 0.72413 | 2.796254 |
| {질병, 스트레스, 가정폭력} => {부정} | 0.001598 | 0.720207 | 2.781076 |
| {성격, 스트레스, 중독, 자살생각} => {부정} | 0.002035 | 0.716599 | 2.767144 |
| {스트레스, 중독, 자살생각, 가정폭력} => {부정} | 0.001218 | 0.716216 | 2.765665 |
| {스트레스, 중독, 자살생각} => {부정} | 0.002736 | 0.706231 | 2.727109 |
| {질병, 스트레스, 중독, 자살생각} => {부정} | 0.001437 | 0.706214 | 2.727044 |
| {성격, 질병, 스트레스, 중독} => {부정} | 0.0012649 | 0.705128 | 2.722848 |
| {성격, 스트레스, 중독, 가정폭력} => {부정} | 0.0010464 | 0.700000 | 2.703046 |
| {질병, 자살생각, 학교폭력, 가정폭력} => {부정} | 0.0011154 | 0.697841 | 2.694712 |
| {스트레스, 중독, 가정폭력} => {부정} | 0.0013454 | 0.692307 | 2.673342 |
| {질병, 학교폭력, 가정폭력} => {부정} | 0.0013569 | 0.690058 | 2.664657 |
| {질병, 성} => {긍정} | 0.0056579 | 0.410684 | 2.654961 |
| {스트레스, 자살생각, 가정폭력} => {부정} | 0.0033349 | 0.687203 | 2.653633 |
| {성격, 스트레스, 자살생각, 가정폭력} => {부정} | 0.0018514 | 0.676470 | 2.612187 |
| {스트레스, 중독, 자살생각, 학교폭력} => {부정} | 0.0011729 | 0.671052 | 2.591266 |
| {성격, 스트레스, 중독} => {부정} | 0.0022887 | 0.670033 | 2.587331 |

4. 요약 및 결론

- 본 연구의 목적은 소셜 빅데이터를 활용하여 데이터마이닝의 연관규칙과 의사결정나무 분석을 통해 우리나라 청소년의 우울 동향과 위험요인을 분석하는 것임
- 우울관련 온라인 문서는 개학시기인 3월부터 증가하여 5월에 감소하며, 다시 여름방학 이후 9월부터 증가하여 11월부터 감소 추세를 보이며, 월요일, 화요일 수요일에 높은 추이를 보이는 반면, 주말에는 감소하는 것으로 나타남
- 우울 관련 부정적인 감정(위험)을 나타내는 문서는 2012년에는 대구광역시, 2013년에는 광주광역시, 2014년에는 울산광역시에서 가장 높았음
- 우울관련 부정적인 감정(위험)을 나타내는 문서는 2014년 24.5%로 나타나 청소년건강행태 온라인조사(2014, 교육부 등)의 우울증상 경험률(26.7%)과 비슷한 결과를 보였음
- 우울 원인에 대한 연관성 예측에서 온라인 문서에 ‘질병’과 ‘가정폭력’이 동시에 언급되면 언급되지 않는 문서보다 우울 감정이 부정적(위험)일 확률이 높은 것으로 조사되었음
- 청소년은 온라인상에서 자살생각과 같은 우울한 담론을 주고받으며, 이러한 언급이 실제적인 자살과 관련된 심리적 · 행동적 특성으로 노출될 수 있기 때문에, 온라인상에서 청소년 우울의 위험 징후가 예측되면 실시간으로 개입할 수 있는 정부차원의 우울관리 모니터링 체계가 구축되어야 함
- 청소년은 주로 학기 중에 우울한 감정 노출이 높게 나타나기 때문에 학기 중 정기적으로 스트레스를 해소할 수 있는 정신건강 관련 교육을 실시할 필요가 있음
- 사이버 상에서 우울한 감정을 주고받는 위기청소년에 대한 대응을 위해서 위기 청소년에 대한 위기 관리 정보제공은 물론 학부모와 교사를 대상으로 위기관리 정보제공과 함께 관련 교육 프로그램이 마련되어야 함

집필: 송태민(정보통계연구실 연구위원) 문의: 044-287-8201

발행인 | 김상호 발행처 | 한국보건사회연구원

(30147)세종시 시청대로 370 세종국책연구단지 사회정책동(1층~5층) | TEL 044)287-8000 | FAX 044)287-8052 | <http://www.kihasa.re.kr>

한국보건사회연구원 홈페이지의 발간자료에서 온라인으로도 이용하실 수 있습니다. <http://www.kihasa.re.kr/html/jsp/publication/periodical/focus/list.jsp>