안녕하세요, 저희는 커뮤니케이션 텍스트마이닝 기말 프로젝트 8조입니다.

저희는 공인의 논란 인정 및 대중의 여론 변화 분석을 주제로 프로젝트를 진행하였습니다.

목차는 다음과 같습니다.

먼저 연구 배경과 가설, 다음으로 데이터 수집 및 전처리 분석 기법을 알아 보고 이를 통해 사례별 분석 결과와 결론을 제시하며 발표를 마치겠습니다.

1.

여러분 혹시 이 광고를 보신 적이 있으신가요? 이는 한때 의류 기업 무신사에서 올린 공식 광고인데, 이를 올리자마자 박종철 고문치사 사건의 상황을 희화한 멘트를 사용했다는 비난이 거세게 일어나 기업의 이미지에 큰 타격을 주었던 광고입니다. 그러나, 지금의 무신사를 보면 알 수 있듯이, 당시 일으킨 파장에 비해 기업의 가치에는 큰 타격이 없었는데요. 이는 당일 빠른 조치와 여러번 거듭한 진정성 있는 사과 덕분이었습니다.

이처럼 일반적으로 기업의 경우, 위기관리 대응책으로는 신속하고 진정성 있는 사과가 효과적이라고 알려져 있습니다. 저희는 이에 대해, 기업이 아닌 사람의 경우에도 이러한 전략이 유효할지에 대한 궁금증이 생겨 연구를 진행하게 되었습니다.

우선, 기업의 위기관리에 대해서는 앞서 언급한 바와 같이 사과의 유용성이 알려져 있습니다.

추가로, 공인의 위기 관리에 대한 선행 연구를 찾아본 결과, 연예인의 경우 논란의 유형과 사과의 주체가 중요하다는 관점의 연구가 있었습니다. 저희는 이에 대해 기업에서처럼 대응의 시점이나, 내용 또한 중요한 변인일 거라는 생각이 들어 이를 바탕으로 가설을 세우고 연구를 진행하였습니다.

이런 두 선행 연구를 기반으로, 저희는 실제 사례들을 통해 사과의 시점과 사과문 내 대응 전략에 따른 여론의 반응 변화를 분석하고자 합니다.

저희는 위기 상황을 그들의 명성에 부정적 영향을 미치는 사건으로 규정하였고, 기업 사례에서 사과가 신뢰 회복의 중요한 수단이었던 것처럼, 공인 역시 위기 상황에서 적절한 사과가 여론을 긍정적으로 전환하는 전략이 될 수 있는지를 검증하는 것이 주요 목적입니다.

이를 검증하기 위해 저희는 사건 발생 이전부터 사과문 게재 이후까지의 기사, 영상, 댓글 데이터를 수집해 여론의 집중 시기와 사과 시점을 비교해 대응 속도를 판단했고, 질적 연구로 사과문 내 전략을 구분해 여론 변화 양상을 분석하였습니다

2.

저희는 프로젝트에 있어 다음 질문에 초점을 두고, 가설을 세워 연구를 진행하였습니다.

첫 번째로, 사건 대응 시점이 여론 변화에 미치는 영향은 무엇인가?라는 질문과 이에 대해 사건 대응 시점이 빠를수록 여론이 긍정적으로 전환될 가능성이 높다는 가설을 세웠고.

둘째, 여론 변화에 영향을 미치는 추가적 요인은 무엇인가?라는 질문에는 평소 이미지가 좋을 수록 부정적 여론이 더빠르게 완화되는지를 알아보고자 하였고

마지막으로, 사과문 내 전략, 즉 책임 인정, 부분 인정, 부인 및 회피에 따라 여론은 어떻게 달라지는가? 라는 질문에 사과문 내의 대응 방식에 따라 여론의 변화 양상이 달라질 것이다라는 가설을 세웠습니다.

이 세 가설에 기반해 저희는 분석을 진행하였습니다.

3.

저희는 사건 전후와 대응 전후의 여론의 변화를 파악하기 위해, 이를 기준으로 시간을 3구간으로 나누었습니다. 데이터는 총 4종류, 네이버 기사, 댓글, 유튜브 영상 및 유튜브 댓글을 수집하였습니다 .

우선 유튜브의 경우, 데이터가 워낙 많기에 여론을 반영할 수 있게 transcript와 각 시간 구간별로 조회수가 높은 10개의 영상, 그리고 각 영상에 대해 좋아요 개수가 10개 이상인 댓글들만 선별적으로, API를 통해 수집하였습니다.

네이버 기사의 경우, request 모듈과 Beautifulsoup을 활용하여 URL을 수집한 후 url내의 기사들을 수집하였습니다. 네이버 기사는 더 많은 기사를 수집하기 위해, 기간을 위기발생 - 1달 ~ 위기 대응 + 1달로 설정하여 모두 수집하였습니다. 이때, 기사 댓글의 경우 실시간 데이터라 request가 아닌 selenium을 활용하여 수집하였습니다.

다음으로, 이렇게 다양한 방식으로 수집한 데이터들을 한번에 일률적으로 처리하기 위해 저희는 다양한 시도를 통해 데이터를 일률적으로 처리하고자 노력하였습니다. 우선, 데이터를 text 형태가 아닌 csv형태로 수집하여 병합한 후, date time object로 변환하여, 각 시간 구간 및 변화를 한번에 처리할 수 있도록 형태를 변환했고, global함수를 통한 변수 자동화와 text dictionary를 활용하여 구조적인 접근이 가능하도록 하여 다양한 분석이 가능하도록 정제하였습니다.

다음으로는 텍스트 분석을 위한 전처리를 비롯해, “삭제된 댓글입니다” 와 같은 불용어를 제거하였고, 위 과정을 거친 데이터로 분석을 진행했습니다

4.

다음은 저희가 이번 연구에 프로젝트에 사용한 데이터 분석 기법에 관한 설명입니다. 우선 빈도 분석과 이를 시각화 한 워드클라우드를 통해 각 인물에 대한 중점적인 반응의 변화를 파악하고자 하였습니다.

여기서, 빈도 분석이란 텍스트 데이터에서 단어의 출현 빈도를 분석하는 것으로 주요 키워드가 무엇인지 파악하고, 시간에 따라 단어의 등장 패턴이 어떻게 변하는 지를 확인하는 기법이며, 해당 결과를 출현 빈도를 바탕으로 키워드를 한눈에 볼 수 있도록 시각화 하는 것이 워드 클라우드입니다. 이 과정을 통해 논란 전, 논란 이후부터 사과문 이전, 사과문 이후 등 기간별로 단어 출현 빈도를 비교할 예정입니다.

빈도분석 코드를 간단하게 살펴보자면, 우선 빈도분석을 하기 위해 간단한 한글 텍스트 전처리 과정을 거친 뒤에 빈도가 가장 높은 단어 30개를 추출했습니다. 그리고 이를 워드 클라우드로 시각화하기 위해 각 사례 및 매체별 워드 클라우드를 생성했습니다.

다음으로 저희가 사용한 데이터 분석 기법은 ‘BERT 감성분석과 여론 추세선 그리기’입니다. BERT란 구글에서 개발한 사전 훈련된 언어 모델로, 문맥을 양방향에서 이해하여 텍스트의 의미를 파악하는 데 뛰어난 모델입니다. BERT 감성분석 코드부터 살펴보면 BERT는 문장 단위로 쪼개서 데이터를 처리해야 하기에 네이버 뉴스의 경우 sent\_tokenize 함수로, 유튜브의 경우 split 함수를 통해 구분자로 쪼개준 후, 이렇게 쪼개진 데이터를 KOTE 모델을 통해 분석을 돌려주었습니다.

KOTE 모델에 대해 좀 더 자세히 설명하자면, 이는 HUGGING FACE TRANSFORMERS 라이브러리에서 제공하는 BERT 기반 감정 분석 모델로, 다양한 플랫폼에서 수집한 50,000개의 댓글에 44개 정서로 레이블링한 데이터셋을 가지고 있습니다. KOTE 모델은 다양한 사이트에서 응용될 수 있고 감정 분석에 특화되어 있기에 이번 프로젝트에 사용하기로 했습니다. 그리고 분석 과정을 이어서 설명하자면 앞에서 네이버 뉴스와 유튜브 내용을 문장 단위로 쪼개준 데이터를 pretrained 모델에 넣고 돌려주고 나온 결과를 긍정, 중립, 부정 감정별로 각각 1,0,-1로 라벨링 했습니다.

다음으로는 여론 변화 추세선을 그리는 코드입니다. 추세선을 그리는 이유는 특정 사건이 발생했을 때 개별 데이터 포인트만 보아서는 흐름이 보이지 않기 때문에 긍정, 중립, 부정 감정이 시간에 따라 어떻게 변화하는지를 시각화하면 사건의 영향을 더 명확하게 보여주기 위한 과정입니다.

먼저 전처리 및 필터링 단계에서 감성 점수 데이터를 날짜 단위로 평균화하고, 각 사건의 영향을 집중적으로 분석하기 위해 기간을 사건 전후 60일로 제한했습니다.

다음으로 시각화 단계에서는 산점도를 활용해 원본 데이터를 표현했습니다. 하지만 데이터가 많을 경우 추세가 흐려질 수 있는 단점이 있어, 데이터 포인트 개수를 최대 30개로 제한하고 무작위로 표본을 추출했습니다. 이를 바탕으로 이동평균법을 적용해 감성 점수의 변화를 평활화하여 전체적인 추세선을 시각화했습니다.

마지막으로 사건에 따른 감정 변화를 강조하기 위해 주요 사건 날짜를 그래프에 수직선으로 표기했습니다.

5.

데이터 분석 결과 저희는 여론이 변화하는 양상에 따라 여론이 좋아진 경우, 악화된 경우, 유의미하지 않은 변화가 나타난 경우 이렇게 세가지로 나누어 분석 결과를 제시하였습니다. 여론이 긍정적으로 변화한 경우로는 설민석, 설현, 박나래의 사례가 있었습니다. 이 세 사례는 모두 감정 추세 양상이 비슷한 모습을 보였는데 우선 설민석의 사례를 살펴 보면, 설민석은 2020년 12월 29일 석사 논문 표절 논란이 있었고, 논란이 발생한 당일에 바로 사과를 했었습니다. 추세 그래프를 보시면 논란 발생과 사과문 기재와 동시에 긍정 추세가 증가하고 부정 추세가 하락하는 모습입니다. 추세 그래프 옆에는 빈도분석을 바탕으로 한 사과문 이후의 워드 클라우드입니다.

설민석과 같이 긍정적 여론 형성 카테고리에 속한 나머지 인물인 설현과 박나래를 살펴보겠습니다.

설현은 2016년5월3일 역사 인식 말실수 논란이 있었고 5월 12일에 사과문을 올렸습니다.

박나래의 경우 2021년3월24일 성희롱 발언 논란이 있었고, 바로 다음날인 3월 25일에 사과문을 올린 뒤 4월9일에 복귀했습니다. 두 사례 모두 설민석과 유사하게 감정 추세 그래프가 사과문 기재 이후 긍정적 여론이 형성되었음을 알 수 있습니다.

다음으로는 빈도 분석 결과표를 보시면서 사례별로 좀 더 자세히 살펴보겠습니다. 표의 시기1,2,3의 경우 각각 논란 이전, 논란 발생과 사과문 기재 사이, 사과문 기재 이후입니다. 우선 설민석의 경우 논란 이전에는 논란 발생 시점을 기사화된 시점으로 정했기 때문에 최초 논란 발생 시점이 명확하지 않아서 그런지 관련 부정 단어가 다수 출현했습니다. 그리고 이후를 보면 논문 표절 관련 단어와 잘못 등의 부정 단어가 출현했으며, 설민석의 논란에 대한 빠른 인정과 대응 관련 단어도 등장하곤 했습니다.

설현의 경우, 논란 이전에는 긍정 단어가 거의 없었으며, 사건 발생 이후에는 사건 관련 인물과 부정 단어가 출현했고, 사과문 기재 이후에는 관련 단어와 사과 등 부정 단어가 다수 등장했습니다.

박나래의 경우, 시기1에는 박나래에 대해 재밌다는 긍정 평가가 존재했고, 시기 2에는 사건에 대한 직접적인 언급 및 사과, 불편 등 부정 단어가 출현했습니다. 시기3에는 기존의 부정 단어뿐만 아니라 하차 언급 등 부정 단어의 빈도수가 증가했습니다.

다음은 부정적 여론이 형성된 경우입니다. 해당 카테고리에 분류된 사례들은 유희열, 양팡, 홍진영, 강민경, 조현아가 있었고 먼저 유희열의 경우를 보면 2022년 6월 14일 표절 의혹이 제기되었고, 그 날 바로 1차 사과문을 기재했습니다. 그 후 여론이 부정적으로 형성되기 시작했고 2차 사과문 직후 이는 더욱 심화되었습니다. 워드클라우드를 보더라도 표절 관련 단어들과 사과 등의 단어가 보이는 것을 알 수 있습니다.

그 뒤로 비슷한 양상의 추세를 보이는 경우들은 간략하게 설명하고 넘어가자면, 양팡의 경우 2020년 8월 5일 뒷광고 논란이 있었고 그 날에 바로 사과문을 올렸으나 여론은 악화되었습니다.

홍진영은 2020년 11월 5일 표절 의혹이 제기되었고, 그 다음날인 11월 6일에 이를 부인하는 입자을 표명하면서 여론이 나빠졌습니다. 이후 2020년 12월18일 사과문을 올렸음에도 여론이 좋아지지는 않았습니다.

강민경의 경우 2020년 7월 15일 뒷광고 논란이 터졌고, 이틀 뒤인 2020년 7월 17일 사과문을 올렸지만 여론은 부정적으로 변했습니다.

마지막으로 조현아는 2014년 12월 5일 갑질 논란으로 인해 구설수에 올랐고, 이에 대해 12월 8일에 대한항공 측 공식 사과문을 기재하였으나 부정적 여론만 심해졌습니다.

지금부터는 사례별로 시기별 빈도 분석 표를 살펴보겠습니다. 유희열의 사례부터 보시면 시기1에는 노래나 영상에 대한 칭찬 등 긍정 단어가 대부분이었으나, 시기3에서 표절 의혹과 이에 대한 입장 및 사과 요구 등의 부정 단어가 다수 등장했습니다.

양팡의 경우 시기1에서는 양팡의 방송 관련 긍정 단어가 출현했으나, 시기 2에서 뒷광고 관련 단어와 사과, 잘못, 문제 등의 부정 단어가 등장했으며 이는 시기3에서도 여전히 존재했습니다.

홍진영을 보면 논란 이전에는 홍진영 관련 단어가 등장하긴 하지만 긍정 및 부정 단어는 거의 없었습니다. 이후 시기3에서 표절 의혹 관련 단어 및 이에 대한 부정 단어가 다수 출현했습니다.

다음 강민경 사례를 살펴보면 뒷광고 논란 이전에는 강민경에 대한 칭찬 관련 긍정 단어가 등장했었습니다. 그리고 시기2에서는 뒷광고 관련 단어와 강민경에 대한 실망 및 해명 요구 단어가 출현했고, 시기3에서 해당 사건에 대한사과 요구 등 부정 단어가 재차 등장했습니다.

마지막으로 조현아의 경우, 시기1에서 조현아 전 대한항공 부사장을 지칭하는 단어가 등장했고, 긍정 단어의 빈도가 많진 않지만 어느 정도는 나왔습니다. 이후 시기3에서 대한항공 관련 단어의 빈도가 급상승했으며 논란 관련 사과와 검찰 조사 관련 단어가 많이 나왔습니다.

마지막 유형으로는 유의미하지 않은 여론 변화를 보인 경우입니다. 해당 유형에 속한 사례는 아이린의 사례 하나밖에 없었고, 이를 자세히 살펴보시면 아이린은 2020년 10월 20일 갑질 의혹이 제기되었고, 2020년 10월 22일에 이에 대한 사과문을 기재했습니다. 워드클라우드 역시 아이린에 대한 부정적인 단어가 눈에 띄는 모습입니다.

빈도 분석 결과를 보시면 의혹 이전에는 아이린의 평소 이미지나 외모와 관련해서 긍정 단어가 출현했고, 시기3에서는 인성, 사과, 잘못 등 논란 관련 부정 단어가 등장했습니다.

다음으로는 시점별 여론 집중도 그래프를 보자면, 위 그래프는 사례별로 시기별 댓글 개수를 카운트한 그래프입니다. 추세선과 마찬가지로 그래프상의 세로 점선은 논란 발생 시기와 사과문 게재 시기를 의미합니다. 대체로 여론 집중도와 대응 시기는 유사한 시기에 맞물린 걸 볼 수 있지만, 그중에서 가장 빠르게 반응한 케이스를 보면 설민석의 경우인데, 초록색 네모로 표시한 부분입니다. 설민석은 논란 발생과 사과문 게재 일자가 일치한 특수 케이스이고, 게재 이후 관심도가 하락한 걸 볼 수 있습니다. 빨간색 네모로 체크한 세 경우는 설현, 유희열 홍진영의 케이스인데, 이중 유희열과 홍진영은 각각 회피와 부인 후 사과한 케이스이고, 셋 모두 사과한 시점에 급격한 관심도가 몰린 것을 볼 수 있습니다. 즉 사과를 미룬다고 관심도가 떨어지는 것은 아니라는 인사이트를 추가로 얻을 수 있고, 앞선 추세선 분석과 해당 그래프를 연결짓자면, 빠른 사과가 긍정적 여론 변화를 유발하진 않지만, 느린 사과의 경우에는 다소 부정적인 효과를 불러일으킨 데에 영향을 끼쳤다고 볼 수 있습니다.

6. 연구에서 저희는 공인의 위기 대응 전략이 여론 변화에 미치는 영향을 분석하기 위해 세 가지 가설을 설정하고 검증하였습니다. 첫째 “사건 대응 시점이 빠를수록 여론이 긍정적으로 전환될 가능성이 높다”는 가설에 대해서는 빠른 사과가 여론 변화에 긍정적 영향을 미친다는 결론에 도달하지는 못했으나, 늦은 사과는 여론에 부정적인 영향을 미쳤음을 알 수 있었습니다. 이를 뒷받침해주는 사례가 홍진영과 조현아의 경우였습니다. 정리하자면 대응 시점이 여론 악화를 막는 데 중요한 역할을 하지만, 단순히 빠른 사과만으로 긍정적 여론을 형성하는 것을 기대하기 어려운 것을 보여줍니다.

둘째, “평소 이미지가 좋을수록 부정적인 여론이 더 빠르게 완화된다”라는 가설은 기각되었습니다. 우선 대상 공인들의 평소 이미지라는 것을 데이터 분석을 통해 수치화하기도 어려웠습니다. 그렇기에 공인들의 평소 이미지에 대해 질적 연구를 통한 사례 조사로 정리해서 가설 검증을 시도했으나 분석 결과 평소 이미지가 긍정적이더라도 위기 상황에서는 동일하게 비판의 대상이 되었고, 여론 회복 속도에 유의미한 차이를 보이지 않았습니다.

셋째, “사과문 내 대응 방식에 따라 여론 변화 양상이 달라질 것이다” 라는 가설 역시 검증 과정에서 사과문 내 대응 방식에 대한 수치화가 힘들었기에 질적 연구로 접근했습니다. 이 가설의 경우 부분적으로 검토되었으나, 사과문에서 책임 회피나 부인이 포함된 경우 여론이 명확하게 악화되었습니다. 하지만 명확하게 책임을 인정한 경우에도 여론이 긍정적으로 전환되지는 않았습니다. 이는 여론의 변화가 사과문 내 대응 전략만이 아닌 사건의 성격과 대중의 반응 등 다양한 요인에 의해 결정됨을 시사합니다.

결론적으로, 대부분의 사례가 윤리적 문제였기에 대중의 수용도가 낮았으며, 빠른 사과가 여론 악화를 막는데 중요한 역할을 한다는 점을 확인할 수 있었습니다. 또한, 공인은 단순히 사과를 넘어서 명확하게 잘못을 인정하고, 꾸준히 반성하는 모습을 보이는 것이 대중과의 신뢰를 회복하는 데 필수적임을 확인하였습니다. 이는 연예인과 같은 공인의 효과적인 위기관리 전략으로서 중요한 시사점을 제공합니다.

잠시 결론이 시사하는 바에 대해 얘기해 보겠습니다

사과는 여론 변화의 만능 열쇠는 아니지만, 신속하고 진정성 있는 사과는 위기 확산을 방지하고 신뢰 회복의 출발점이 될 수 있습니다. 반면, 책임을 회피하거나 논란을 축소하려는 태도는 오히려 부정적 여론을 악화시키고, 장기적으로 이미지 회복을 더욱 어렵게 만듭니다.

따라서 사과의 시점과 내용에 있어 신중하면서도 빠른 대응이 필요하며, 대중과의 공감대를 형성하는 진정성 있는 행동이 요구됩니다. 이번 연구 사례에서도 확인했듯이, 논란의 본질을 회피하지 않고 책임을 명확히 인정하는 태도가 위기 관리의 핵심입니다.

이 교훈은 연예인뿐만 아니라 사람과 관계를 맺는 모든 이에게 적용될 수 있는 기본 원칙입니다. 실수를 저질렀을 때 문제를 회피하거나 애매한 태도를 보이는 것이 아니라, 주체를 명확히 하고 책임을 다하는 것이 신뢰를 회복하는 가장 확실한 방법입니다. 또한, 잘못을 인정하고 진정성 있는 사과를 전하는 과정은 단순히 위기에서 벗어나기 위한 수단이 아니라, 더 나아가 상대와의 관계를 재정립하고 더욱 강한 신뢰를 쌓는 기회가 될 수 있습니다.

마지막으로, 실수 이후의 행동도 중요합니다. 사과와 책임 인정이 끝이 아니라, 이후에도 꾸준한 반성과 개선된 모습을 보여주는 것이 진정한 신뢰 회복의 시작입니다. 공인뿐만 아니라 우리 모두가 이러한 태도를 실천할 때, 개인과 공동체 모두가 더욱 건강하고 신뢰할 수 있는 관계를 만들어 나갈 수 있을 것입니다.

7.

그 다음으로 이번 연구에서 드러난 한계점을 말씀드리겠습니다  
첫 번째는 데이터 오분류입니다. 데이터 수집에 있어서 동명이인이나 키워드 불일치 문제가 발생했습니다. 저희가 분석하고자 시도한 사례 중에서는 배우 지수의 학폭 논란도 포함되어 있었지만, 크롤링 과정에 있어 배우 지수와 블랙핑크 지수를 구분하는 데 어려움을 겪어 결국 해당 사례는 폐기하는 상황이 발생했습니다  
두 번째는 감성 분석 도구의 한계입니다. 감성 분석 도구가 비꼬는 표현이나 중립적 의견과 같은 미묘한 뉘앙스를 정확히 반영하지 못해 일부 부족함이 나타났습니다.

세 번째는 사례의 제한성입니다. 이번 연구는 일부 연예인 사례만 다루었기 때문에 연구 대상의 범위가 협소했기에 연구 결론을 일반화하기엔 다소 부족할 수 있습니다.

네 번째는 실패한 분석 시도입니다. LDA 토픽 모델링과 클러스터링 분석을 시도했으나 결과 해석이 어려워 해당 분석들은 폐기하였습니다.

마지막으로, 유튜브 데이터의 한계입니다. 조현아 사례의 경우 유튜브 활성화 이전 시기적 배경으로 댓글과 콘텐츠 수가 현저히 적어 네이버 데이터를 중점으로 추세를 파악했습니다.

이러한 한계를 보완하기 위해 향후 연구에서 고려해야 할 제언을 드리겠습니다.

첫 번째는 다양한 플랫폼 활용해 심층적인 결과를 도출하는 것이고, 두 번째는 미묘한 뉘앙스를 반영하기 위해 감성 분석의 정확도를 높이는 것입니다. 다음은 사례를 확장해 연구 결론을 일반화시키는 것과, 데이터 수집 과정에서 동명이인 문제나 키워드 불일치 문제를 해결해 분석을 진행하는 것입니다

프로젝트에 있어서 참고 문헌은 다음과 같습니다

감사합니다