|  |
| --- |
| **1. 주제**  ~~유튜브 광고성 댓글 필터링 프로그램~~  **분반, 팀, 학번, 이름**  나 분반,6팀,20243274,남동우 |

|  |  |
| --- | --- |
| **2. 요약**  본 프로그램은 유튜브 댓글에서 음란성 광고 성향의 댓글을 자동으로 작성하는 매크로 봇을 탐지하는 것을 목표로 한다. 프로그램은 자연어 처리(NLP) 기술과 패턴 인식을 활용하여 비정상적인 댓글 활동을 감지하고, 해당 계정을 자동으로 차단하거나 경고하는 시스템을 구현할 예정이다. 이를 통해 유튜브 커뮤니티의 건전성을 유지하고, 사용자 경험을 향상시키고자 하는 것이 목표이다. | **3. 대표 그림**  개발배경: 유튜브에서 일어나는 음란성 광고 매크로봇 차단    그림1. 매크로봇 댓글 예시 |

|  |
| --- |
| **4. 서론**  **1) 배경설명**  최근 유튜브와 같은 대형 소셜 미디어 플랫폼에서는 자동화된 매크로 봇을 이용해 대량으로 댓글을 작성하고, 특히 음란성 광고를 게시하는 사례가 늘어나고 있다. 이러한 봇은 사람의 개입 없이 반복적으로 광고 댓글을 생성하고, 유해한 콘텐츠가 유포될 위험성을 가중시키고 있다. 이러한 문제는 유튜브 커뮤니티의 사용자 경험을 저해하고, 플랫폼의 신뢰성에 부정적인 영향을 미친다.  **2) 문제 정의**  유튜브는 댓글을 통해 사용자 간의 소통을 지원하는 중요한 기능을 제공한다. 하지만 봇을 통한 음란성 광고는 정상적인 사용자 경험을 방해하고, 콘텐츠의 질을 떨어뜨리고 있다. 2020년 유튜브에서 다수의 사용자가 음란성 광고 댓글로 불편을 겪었다는 보고가 있었다. 이는 특정 키워드를 포함한 봇들이 대규모로 동작하면서 사용자의 피드에 불쾌한 콘텐츠가 노출되는 문제였다.또한, 트위터나 인스타그램 등 다른 소셜 미디어에서도 유사한 봇 활동이 발견되고 있으며, 기존의 단순 스팸 필터링으로는 이들을 효과적으로 차단하기 어렵다.이러한 봇의 탐지와 차단이 어려운 이유는 이들이 사람처럼 보이도록 다양한 패턴의 댓글을 남기기 때문이다.  **3) 극복 방안**  본 프로그램은 자연어 처리(NLP)와 패턴 인식 기술을 도입하여 봇이 자동으로 작성하는 음란성 광고 댓글을 탐지하고 차단하는 시스템을 개발하는 것을 목표로 한다. 이를 통해 유튜브 커뮤니티의 건전성을 유지하며, 사용자 경험을 향상시킬 수 있다. |

|  |
| --- |
| **5. 본론**  **1) 시스템 개요**    **2) 필요한 기술 요소**  (1) 자연어 처리(NLP): NLP는 텍스트 데이터를 분석하고 의미를 파악하는 기술로, 음란성 또는 광고성 댓글을 감지하는 데 핵심적인 역할을 한다. BERT와 같은 사전 학습된 모델을 사용하여 텍스트 분류를 수행하고, 댓글의 맥락을 이해하는 능력을 향상시킬 수 있다.  (2) 패턴 인식 및 머신러닝: 봇 활동은 반복적이며 일정한 패턴을 따르므로, 이를 탐지하기 위한 머신러닝 알고리즘을 사용할 수 있다. 예를 들어, KNN 또는 SVM과 같은 알고리즘을 사용해 비정상적인 댓글 패턴을 인식하고, 일반 사용자와 차별화할 수 있다.  (3) 유튜브 API: 유튜브 API를 통해 댓글 데이터를 수집하고 관리할 수 있다. 이를 통해 실시간으로 댓글을 분석하고 봇 댓글을 탐지하는 시스템을 구현할 수 있다.  (4) 데이터베이스: MongoDB 또는 MySQL과 같은 데이터베이스를 이용해 봇으로 의심되는 댓글 데이터를 저장하고, 시간이 지남에 따라 데이터 기반으로 시스템을 개선할 수 있다.  **3) 구현 방법 및 개발 방향**  (1) 데이터 수집:  유튜브 API를 사용하여 실제 유튜브 댓글 데이터 수집 후 광고성 댓글과 일반 댓글을 분류한다. 이는 모델 학습에 이용할 데이터이다.  (2) 모델 학습 및 구축:  수집된 데이터를 바탕으로 NLP 모델을 학습시킨다. BERT와 같은 사전 학습된 모델을 통해 댓글 텍스트를 분석, 분류 모델을 학습시켜 음란성 광고 댓글을 탐지한다.  (3) 패턴 인식 시스템 구축:  머신러닝 알고리즘을 도입해 비정상적인 댓글 작성 패턴을 탐지한다. 일정 시간 내에 유사한 댓글이 반복적으로 작성되거나, 비정상적으로 빠른 시간 내에 여러 개의 댓글을 작성하는 계정은 봇으로 분류한다.  (4) 실시간 댓글 감지 및 차단 시스템 개발 개발된 시스템을 실시간으로 작동하게 하여, 유튜브에 게시되는 댓글을 자동으로 분석하고 봇으로 의심되는 댓글을 즉시 차단하거나 관리자에게 경고를 보낸다. |

|  |
| --- |
| **6. 결론**  **1) 내용 요약**  본 프로그램은 유튜브의 음란성 광고 댓글 매크로 봇을 탐지하고 차단하는 시스템을 구축하는 것을 목표로 합니다. 자연어 처리와 머신러닝 알고리즘을 이용해 음란성 또는 광고성 댓글을 탐지하고, 봇이 작성한 비정상적인 댓글 패턴을 파악하여 차단합니다. 유튜브 API를 통해 실시간으로 데이터를 수집하고 분석하며, 이를 바탕으로 봇 탐지 시스템을 지속적으로 개선해 나갈 것입니다.  **2) 향후 할 일**  음란성 광고 댓글의 다양한 패턴을 탐지하기 위한 추가 데이터 수집 및 모델 학습을 진행하고, 실시간으로 봇 활동을 감지하고 차단하는 시스템의 성능을 최척화 시키기 위한 알고리즘 연구를 진행한다. 또한 이 시스템을 인스타그램과 같은 다른 소셜 미디어 플랫폼에서도 적용할 수 있도록 확장 가능성을 검토해 본다. |

**7. 출처**

[1] 홍광수, 한재근 **End-to-End BERT: 만능 모델 BERT 학습부터 추론** NVIDIA Korea 2020.02.20

https://blogs.nvidia.co.kr/blog/bert/