problem description, data, methodology/algorithm, related work and evaluation plan

1. 유튜브 댓글 분석(긍정도 점수)
   1. Data : 스트리밍 사이트(유튜브 ..), 영화 리뷰. 앱 리뷰?
   2. Algo : 주요 단어, 맥락 등 평가의 긍정/부정 출력 - 가중치, 빈도수 등등
   3. Evaluation : 별점 데이터로 채점, 좋아요/싫어요 비율(?)등과 비교,
2. Rhyme 맞는 가사 generator
   1. Data : syllable 데이터, 사전 데이터, 가사를 만들기 위한 corpus 들
   2. Algo :
      1. 영어 - 수업시간에 배운 syllable?음절단위로 발음기호 적어주는 것 이용해서 유사한 단어 찾기  
         한국어 - 초성/중성/종성 같은거 찾기
      2. 찾은 단어가 들어있는 문장을 corpus에서 찾아서 그 문장을 사용한 가사 만들기
   3. Evaluation :
      1. 직접 불러보기…..(최다은이)ㅋㅋㅋㅋㅋㅋㅋㅋㅋㅋ->최다은 222
      2. 뭐가 있을까?
3. 또는 작사 작업을 할 때 원하는 개수의 음절을 가진 원하는 뜻의 단어 추천?
4. 사전에 없는 새로운 용어(?)/들 추가
5. 나이대 별 언어 사용 파악->어느 시대의 news 인지
   1. Data : 시대별 뉴스 데이터 (인터넷 긁어오기)
   2. Algo : 동사/고유명사 등 시대별 겹치지 않는 표현들을 비교하기
   3. Evaluation : 실제 년도와 비교 /

(news 단어들을 분석해서 그 시대의 trend 파악하기->특정 분야로 한정시킬 수도 있음 (ex 2019 패션계의 trend는?)

1. 누가 쓴 글인지 판별하는 판독기 (유사도 검사 등) ->표절 판독?
   1. Data : 뉴스 기사 데이터, 카톡 데이터?,
   2. Algo : 문장 구조, 혹은 글에서 흔하게 보이는 어려운(?) 단어들 비교하기
   3. Evaluation : 뉴스 기자별로 유사도 출력

(~ 글 A 와 ~ 글 B 가 있을 때 같은 사람이 적은 글일 확률 71% 이런 느낌)

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | 1 | 2 | 3 | 4 |
| 다은 | 이거 | 이거 |  |  |
| 진원 |  | o | o | o |
| 승호 | 0 |  | 0 | 0 |
| 형조 | 0 |  | 0 |  |
| 재현 | 0 | 0 |  | 0 |