

안녕하십니까 3월 30일 비바프로팀의 발표를 시작하도록 하겠습니다.

우선 목차에는 말뭉치 추가에 관한 내용, 웹서비스 진행 현황, 웹서비스 진행 계획 순으로 발표하겠습니다.

저번 기업 미팅 때 유독 F1 Score가 낮았던 PT에 대해서 말뭉치를 추가하였다고 발표 했었습니다.

당시 PT 식물 태그가 1097개에서 9603개로 추가되었고, F1 score가 0.62에서 0.91로 꽤 큰 폭으로 개선된 것을 확인했습니다.

그 뒤로 TR 이론, FD 학문 분야 태그에 대해서 말뭉치 추가를 진행하고자 하였고, 현재 결과를 기다리는 중에 있습니다.

이후 결과가 괜찮다면, 그 외에도 추가하여 진행할 예정입니다.

웹서비스 진행 현황입니다. 먼저 간단하게 기능적인 부분을 테스트하기 위해 로컬에서 진행하였습니다.

터미널에서 서버와 클라이언트 부분을 실행한 뒤,

서비스 페이지에서 입력 부분에 입력을 한 뒤 저장 버튼을 누르면

정해진 경로에 txt 파일이 저장되고, 저장된 파일을 이용하여 prediction 스크립트가 실행됩니다.

그 때 생성된 결과를 읽어 결과 부분에 출력해 줍니다.

앞으로의 진행 계획들은 크게 세가지 부분으로 전처리와 후처리, 추가기능이 있습니다.

전처리는 여러번 언급했던 것처럼 입력 시 맞춤법 검사를 진행하는 부분입니다.

그러나 사용자가 의도하는 바와 달라질 수 있기 때문에 체크 박스를 추가할 생각입니다.

그래서 체크 박스를 클릭하지 않을 경우는 입력한 그대로 사용되게 만들 예정입니다.

반면에 체크 박스를 클릭하게 되면, 맞춤법 검사를 통해 올바른 입력이 들어가게 할 예정입니다.

후처리는 기업 미팅 때 나왔던 개체명의 단어 부분을 토큰나이징 하기 위하여 soynlp 라는 패키지를 사용할 예정입니다.

해당 패키지를 사용하여 오늘은 과 김치찌개가 에 해당하는 부분을 명사와 조사로 나눠 오늘. 은, 김치찌개. 가 로 나눠 처리할 예정입니다.

추가 기능 부분으로는 마찬가지로 기업 미팅 때 나온 부분을 보완할 생각입니다.

로그인 기능을 현재 구현 중에 있으며, 로그인 후 과거에 진행했던 서비스 기록을 로그로 저장하여 언제든지 접근 할 수 있게 만들 예정입니다.

그리고 각 로그별로 잘못된 개체명이 있을 시 오류 레포트를 진행할 수 있는 게시판을 만들 것입니다.

이것으로 발표를 마치도록 하겠습니다.