IEEE ACCESS 실험 설계

1. 아이디어  
- 챗지피티에게 입력한 메시지가 스팸인지, 햄인지, 판별 불가한지 단답으로 답해달라는 명령을 넣어 처음 세팅을 합니다. 여러 스팸메시지를 차례로 넣고, 그에 따른 챗지피티의 대답을 그대로 모두 긁어와 한 문서에 저장한 후 실제와 비교하려고 합니다.

2. 과정

1) 첫 명령으로 “내가 이제부터 여러 문자를 보여줄 거야. 그걸 보고 스팸 문자인지 아닌지 판별해주면 돼. 대답은 단답으로 해줘. 스팸문자이면 "스팸", 스팸이 아니면 "햄", 혹시 애매하거나 확실하지 않으면 "판별 불가"라고 말해줘.”를 입력합니다.

2) 이후부터 갖고 있는 메시지 데이터를 한 줄씩 읽히게 합니다. (메시지에서 ‘[(KISA:SOL)]202208241115[(KISA)]202208241104[(KISA)]불법대출[(KISA)]’과 같은 부분은 제거하여 넣으면 좋을 것 같습니다.)

3) 챗지피티가 입력한 메시지에 따라 답을 “스팸”, “스팸입니다”, 등으로 할 텐데, 메시지 입력에 따른 모든 답을 메모장 같은 파일에 하나로 저장합니다.  
(저장 시, 모든 답을 한 줄에 저장하지 않고, 각 답변을 새로운 단으로 나눠서 저장되게 하면 좋을 것 같습니다. 아래와 같이…  
스팸  
스팸  
스팸  
햄  
….

추가) 챗지피티가 얼마나 잘 분류했는지 계산해주는 프로그램도 있으면 좋을 것 같습니다. (바쁘시면 안 해도 됩니다!!)

1) 챗지피티의 답변을 저장한 파일을 다시 읽어서

2) “스팸”이라는 단어의 개수, “햄”이라는 단어의 개수, “불가”라는 단어 개수를 세고

3) 스팸이 전체(스팸+햄+불가)에서 몇 퍼센트인지, 햄이 몇 퍼센트인지, 불가가 몇 퍼센트인지 출력합니다.