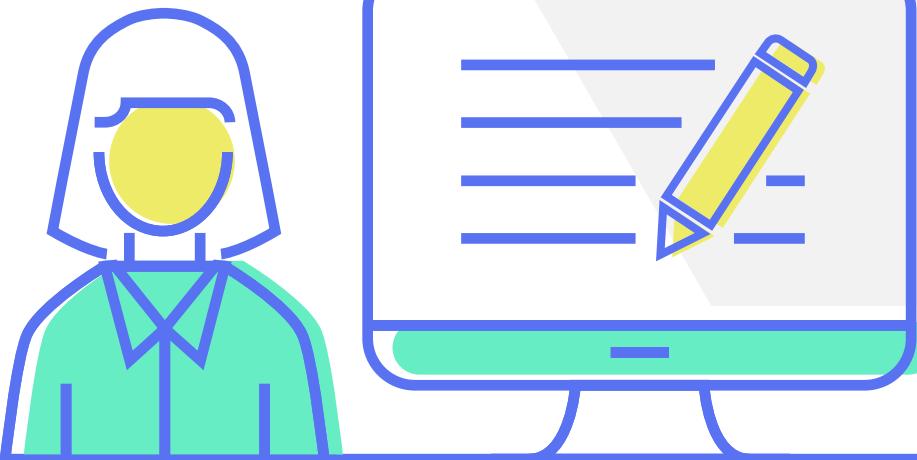


캡스톤 디자인 프로젝트

Conversational Ai

for Education



팀명: Ahtty

201502426 이승준(발표자)
201500429 김민우
201500514 김성진

목차

Contents

01. Feedback

중간 발표 피드백

02. Propose

프로젝트 목표

03. Improve

프로젝트 개선 사항

04. Demo

프로젝트 구현

05. Appendix

별첨

01. Feedback

중간발표피드백

항목	지적 사항	수정 및 보완 사항
주제 관련	1 개발하고자 하는 AI Chatbot이 기존 타사의 제품과 유사점과 차별점을 명확히 기술할 것	기존의 Hugging Face에서 제공하는 Conversational AI 챗봇과 유사점 및 차별점을 데모 비교로 기재하였음.
	2 AFL은 김명숙(2015)의 연구내용을 그대로 적용한 것인지? 일부 보완한 내용이 있는지?	AFL의 연구 내용을 보완하여 본 프로젝트에 적용. 챗봇에서 AFL 기능이 교사의 역할을 대리 수행하면서 학습 상황을 지속해서 평가함. 본 프로젝트에서는 평가 요소로 학습 성취도와 문법 성취도를 평가함.
	3 최종발표 시 줌마영어 회원들을 대상으로 피드백을 받은 내용 을 첨부할 것	최종 발표시 사용자 만족도 조사를 첨부 할 예정.
	4 개인화가 실제로 얼마큼 다른 대화를 이끌어가는지에 대한 검증계획이 있는지?	본 프로젝트에서는 챗봇을 통해 개인화가 아닌 줌마영어 컨텐츠 맞춤화를 진행하였음. 따라서 개인화의 검증 계획은 없음. 대신에 줌마영어 컨텐츠에 맞춤화되어 있는지를 F-1 score, perplexity를 통해 검증하였고 해당 내용을 본문18p에 첨부하였음.
	5 사용자 평가가 실제 사용자의 영어 실력과 상관관계가 있는지 확인할 수 있는가?	현재 줌마영어에서 시행되고 있는 레벨 테스트가 존재하지 않음. 따라서 AFL 기능을 통해 레벨 테스트의 기준을 제공하는 것이 프로젝트의 목표였음. 따라서 상관관계는 확인할 수 없음. 대신에 표본 조사를 통해 사용자의 Acceptance를 반영하여 기준점을 세웠음.

항목		지적 사항	수정 및 보완 사항
모델 성능	6	데모에서 보면 셔츠의 가격을 물어보는데 사이즈에 대한 대답을 하고 있음. 화면을 보면 학습 데이터에 사용된 것과 동일한 질문을 하는데도 잘못된 대답을 하는 이유는 무엇인지?	중간 발표시 줌마영어 교재의 대화데이터만 사용하여 딥러닝 모델이 충분한 학습이 이루어 지지 않았음. 이후, Ai Hub에서 제공하는 영어-한국어 말뭉치를 사용하여 학습량을 증대함
	7	최종 발표 이전에 실제로 사용해 볼 수 있도록 할 수 있는지? 꼭 app 형태가 아니어도 상관없음. 예를 들어 데모에서처럼 학습된 모델을 terminal에서 간단하게 사용해 볼 수 있도록 하면 좋을 것 같음	6월 1주 최종 앱 개발을 마침. 이후 줌마영어 회원들, 전문가 그리고 일반인들에게 앱 사용 후 만족도 조사를 진행중.
사용자 만족도	8	성취도 평가를 위한 AFL에 관한 내용이 clear 하지 않음. 이 chatbot 시스템의 경우, 단순한 grammar checking이나 문장의 유사도를 측정하는 것 외에 chatbot 시스템 내주는 결과에 대한 사용자의 acceptance가 중요할 것 같은데, AFL이 이런 부분까지 포함하고 있는지. 아니면 다른 형태의 검증 방안이 있는지?	MRPC & CoLA 데이터셋을 통해 만들어진 모델로 퍼센트 성적으로 수강생에게 평가를 내어줌. 하지만 퍼센트 결과로는 수강생의 Acceptance를 높이는데 한계점이 존재함. 따라서 수강생의 Acceptance를 높이기 위해 틀린 문장의 Spelling Check 기능을 추가하여 사용자가 이해할 수 있도록 하였음. 또한, 사용자의 Acceptance를 검증하기 위해 줌마영어 수강생들의 앱 만족도 조사를 진행하였음.
	9	프로젝트 목표 및 일정 대비 진척도가 매우 미진함. 핵심 목표 달성을 집중할 필요가 있음	딥러닝 모델 학습 과정에서 생각보다 많은 시간이 소요되었음. 모델 학습을 마친 이후, 앱 개발에서는 빠른 진척도를 보이며 최종 프로젝트를 마무리 지었음.
학습 진척도	10	Chatbot의 목표 수준의 제시가 필요. 구체적인 지표의 정의와 자동 측정, 추후 이를 활용한 개선방안의 보완 등	구체적으로 평가하고자 하는 지표의 정의 및 기준 알고리즘과 비교를 별첨에 첨부하였음.

AFL 사용 목적

AFL Propose

Q. 성취도 평가를 위한 AFL에
관한 내용이 Clear 하지 않음.
사용자의 acceptance 가 중요할
것 같은데, AFL 이 이런 부분까지
포함하고 있는지?

A. AFL이란, 형성평가 위주의 학
습의 결과를 평가하는 전통적인
평가 방식에서 벗어나

**성취기준에 도달하는 과정을
점검하는 평가 방식**으로 아래
3가지 효과를 기대

지속적인 평가



성취과정 점검



피드백의 극대화



AFL 개념의 챗봇 적용

- AFL의 개념에서 착안하여 학습을 위한 평가 방식 고안
- 챗봇이 교사의 역할을 대리 수행
- 챗봇과 지속적인 대화를 통해 수강생의 학습 상황을 지속적으로 평가 및 점검
- 평가를 점수화하여 수강생에게 공개하고 부족한 부분에 대한 동기 부여 제공

AFL 사용목적

AFL Propose

- MRPC & CoLA 데이터셋을 통해 만들어진 모델로 퍼센트 성적으로 수강생에게 평가를 내어 줌
- 퍼센트 결과로는 수강생의 Acceptance를 높이는데 한계점이 존재
- 수강생의 Acceptance를 높이기 위해 틀린 문장의 Spelling Check 기능을 추가

MRPC

Microsoft Research
Paraphrase Corpus

- 컨텐츠를 기반으로 잘 말하고 있는지 유사성 체크
- 퍼센트로 표시

CoLA

Corpus of Linguistic
Acceptability

- 문법적으로 알맞게 대화를 하고 있는지 체크
- 퍼센트로 표시

Bing Spell Check API

Microsoft Azure

- 문법적으로 틀린 부분을 수정하여 수강생에게 전달

02. Propose 프로젝트 목표

기업 요구사항

Enterprise requirements

전화 영어 수업 전 예습 프로그램 필요

- 현재 전화 영어 서비스 ‘딱풀영어’ 시행 중
- 전화 영어 수업 전 교재 내용 예습 프로그램 필요
- 전화 영어 수업 후 효과적인 복습 방법 고안 중

앱 개발 시 음성인식 기능 요구

- 전화 영어 시행 전 과제는 주로 네이버 밴드를 활용
- 줌마영어 회원 대부분이 타이핑에 미숙 → 음성으로 과제 제출

프로젝트 목표

Project Propose

1. 다양한 상황에서 자유로운 대화가 가능한 챗봇
2. 교육내용(컨텐츠)기반 교육 플랫폼



FRAMEWORK

- 컨텐츠를 기반으로 Conversational AI 개발
- 성취도 평가를 위한 AFL시스템 구축
- 최종적으로, Conversational AI와 AFL이 결합된 프로세스 구축

APPLICATION

- 컨텐츠 기반 단어, 문장 학습이 가능한 기본 교육 기능
- 컨텐츠를 기반으로 한 자유대화 플랫폼 제공으로 응용 교육 기능 제공
- 음성인식 기능 활용

03. Improve

프로젝트 개선사항

프로젝트 프로세스

Project Process

Data Collection

대화형 Ai Dataset

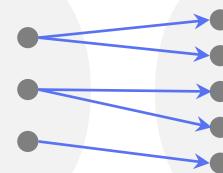
- 줌마 영어 컨텐츠
- AI Hub

AFL Dataset

- CoLA, MRPC

Data Matching

⟨Persona⟩ ⟨상황⟩



Text Preprocessing

1. Expanding Contraction
2. Tokenization
3. Convert to Lower Case

Data Augmentation

Personality Permutation

Persona 1	Persona 3
Persona 2	Persona 4
Persona 3	Persona 1
Persona 4	Persona 2

Model Training

대화형 Ai 모델

- GPT2 Pre-Trained
- AFL
- MRPC : 3시간,
- CoLA : 4시간 학습

Application Dev.

- 대화형 Ai와 AFL모델을 REST API 형태로 구축
- Flutter 사용

Decoder Setting

Hugging Face Decoder
세팅 중 Medium 값 사용

- Top-k : 50%
- Top-p: 80%
- Temperature

AFL

테스트를 통해 각 모델의
임계치 설정

- MRPC : 50%
- CoLA: 80%

각 임계치 초과시 다른
상황 대화 가능

Model Selection

모델 후보군 생성
↓
성능 테스트 후 최적 모
델 선정

Parameter Optimization

데이터셋 서버환경에 맞게 수정

ConvAI2 데이터 구조

ConvAI2 Data Structure

Persona 1	Persona 2
I like to ski	I am an artist
My wife does not like me anymore	I have four children
I have went to Mexico 4 times this year	I recently got a cat
I hate Mexican food	I enjoy walking for exercise
I like to eat cheetos	I love watching Game of Thrones

[PERSON 1:] Hi
[PERSON 2:] Hello ! How are you today ?
[PERSON 1:] I am good thank you , how are you.
[PERSON 2:] Great, thanks ! My children and I were just about to watch Game of Thrones.
[PERSON 1:] Nice ! How old are your children?
[PERSON 2:] I have four that range in age from 10 to 21. You?
[PERSON 1:] I do not have children at the moment.
[PERSON 2:] That just means you get to keep all the popcorn for yourself.
[PERSON 1:] And Cheetos at the moment!
[PERSON 2:] Good choice. Do you watch Game of Thrones?
[PERSON 1:] No, I do not have much time for TV.
[PERSON 2:] I usually spend my time painting; but, I love the show.

- History에서 주어진 문장(홀수 개)을 통해 다음 문장 예측 → 훈련
- Candidate : 다음 발화의 예측에 사용되는 문장, 마지막줄 정답
- ⟨Persona 1⟩과 ⟨Persona 2⟩를 분리하여 Dialog 생성
- Data Augmentation을 위해 Persona 순서를 바꿔가며 History 생성

출처: Thomas Wolf. "TransferTransfo: A Transfer Learning Approach for Neural Network Based Conversational Agents". 3page,
Zhang, S. "Personalizing Dialogue Agents: I have a dog, do you have pets too?". 2page

PERSONA ConvAI2 'Persona 1'

I like to ski.
My wife doesn't like me anymore.
I've went to Mexico 4 times this year.
I hate Mexican food.
I like to eat cheetos.

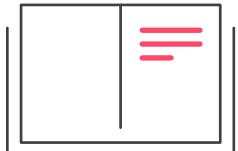
CANDIDATE

Awesome . So what do you do for a living ?
You are telling the truth lol
Haha yea. I love cookies . I sleep way to much to . You like music ?
That is nice to know . I travel due to being financially stable.
I am a karate instructor and love it
...
Great, thanks! My children and I were just about watch Game of Throne.

HISTORY

⟨speaker1⟩ : Hi
⟨speaker2⟩ : Hello! How are you today?
⟨speaker1⟩ : I am good thank you, how are you.
⟨speaker2⟩ : Great, thanks! My children and I were just about watch Game of Throne.

데이터셋 예시



줌마영어(교재)

In a Hospital

〈Doctor〉 : How are you today?
〈Zumma〉 : I'm not feeling well now
〈Doctor〉 : Let's See
〈Zumma〉 : I've got a headache and a fever.
〈Doctor〉 : How Long Have You Been Sick?
〈Zumma〉 : For About 2 Days.
〈Doctor〉 : Do you have a Cough, too?
〈Zumma〉 : Yes, sometimes I do

항목	값
교재 챕터수	19
대화턴	12~14
생성 가능 Persona	19
생성 가능 History	11~13

Persona

I'm not feeling well now.
I have got a headache and a fever.
I have been a headache about 2 days.
Sometimes, I have a cough.



Ai Hub 상황 데이터

증상을 묻고 답하는 상황

〈Speaker1〉 : Hello, nice to meet you, what brings you here?
〈Speaker2〉 : I have fluid coming out of my right ear and I have pain inside the ear canal.
〈Speaker1〉 : Let's have a look. Do you have any fever or sore throat?
〈Speaker2〉 : I've had a little chill and a slight fever for the last few days.



항목	값
상황수	2,769
대화수	100,000
최대 생성 Persona	2,769
최대 생성 History	25,000

History

〈Speaker1〉 : Hello, nice to meet you, what brings you here?
〈Speaker2〉 : I have fluid coming out of my right ear and I have pain inside the ear canal.
〈Speaker1〉 : Let's have a look. Do you have any fever or sore throat?
〈Speaker2〉 : I've had a little chill and a slight fever for the last few days.

데이터셋 매칭

Dataset Matching

- 줌마 영어 교재 챕터 당 여러 상황이 매치됨
- 일대일 관계 → 일대다 변형 필요

항목	값
교재 챕터수	19
대화턴	12~14
생성 가능 Persona	19
생성 가능 History	11~13
〈줌마 영어 교재 데이터셋 요약〉	

항목	값
상황수	2,769
대화수	100,000
최대 생성 Persona	2,769
최대 생성 History	25,000

〈Ai Hub 상황 데이터셋 요약〉

줌마 영어 교재

In a Hospital

- 증상을 묻고 답하는 상황1
- 증상을 묻고 답하는 상황2
- 증상을 묻고 답하는 상황3
- ...

In a Taxi

- 택시 이용
- 택시 찾기
- 택시 요금 계산
- 입국 심사 할 때
- 입국/출국 상황

Immigration

...

Ai Hub 상황 데이터

학습 데이터 보강

Data Augmentation



줌마영어(교재)



Ai Hub 데이터

항목	본 프로젝트	설명
Persona	4	Persona 당 문장 수
Utterance	3	Persona당 대화턴수
History	4	History 당 문장 수
Total	1,718	총 데이터셋

출처: Zhang, S. "Personalizing Dialogue Agents I have a dog, do you have pets too?". 4page

PERSONA 줌마영어 교재 'In a Hospital'

1. I'm not feeling well now.
 2. I have got a headache and a fever.
 3. I have been a headache 2 days.
 4. Sometimes, I have a cough.
-
3. I have been a headache 2 days.
 1. I'm not feeling well now.
 4. Sometimes, I have a cough.
 2. I have got a headache and a fever.

CANDIDATE 다음 문장 예측을 위해 사용 → 임의의 19개 문장 추출

I want to change my cleanser, can you recommend a good one?
There is no shower room, so you should take a shower in your room.
Then can you stop in front of the dry sauna there?
But what if I get any questions, sir?
What do you want to know? I will tell you the parts I know.
...

HISTORY Ai Hub '증상을 묻고 답하는 상황'

⟨Speaker1⟩ : Hello, nice to meet you, what brings you here?
⟨Speaker2⟩ : I have fluid coming out of my right ear and I have pain
inside the ear canal.
⟨Speaker1⟩ : Let's have a look. Do you have any fever or sore throat?
⟨Speaker2⟩ : I've had a little chill and a slight fever for the last days.

텍스트 전처리

Text Pre-Processing

1. Expanding Contractions

I'm not feeling well now.



I'll be staying for three days.

I am not feeling well now.

I will be staying for three days.

2. Tokenization

I am not feeling well now.



I will be staying for three days.

['I', 'am', 'not', 'feeling', 'well', 'now', '.']
['I', 'will', 'be', 'staying', 'for', 'three', 'days', '.']



3. Converting all Characters to Lowercase

['I', 'am', 'not', 'feeling', 'well', 'now', '.']



['I', 'will', 'be', 'staying', 'for', 'three', 'days', '.']

['i', 'am', 'not', 'feeling', 'well', 'now', '.']
['i', 'will', 'be', 'staying', 'for', 'three', 'days', '.']

4. Lemmatization

['I', 'am', 'not', '**feeling**', 'well', 'now', '.']



['I', 'will', 'be', '**staying**', 'for', 'three', '**days**', '.']

['i', 'be', 'not', **feel**', 'well', 'now', '.']
['i', 'will', 'be', **stay**', 'for', 'three', '**day**', '.']

훈련 데이터셋 & 검증 데이터셋 분할

- 훈련 데이터셋 80%, 검증 데이터셋 20% 분할
- 대용량 JSON 포맷으로 출력

딥러닝 PC 환경 구축

- Ubuntu 18.04 LTS
- CUDA 10.1 사용
- GTX 2080Ti
- SSH 서버 사용

모델 학습

Model Training

Open AI GPT-2 vs. B.E.R.T

- 사전 학습 모델 중 NLP 분야에선 BERT 주로 사용
- BERT는 기본 Transformer의 Decoder 사용 X
- 완전한 문장을 기준으로 학습되었기 때문에 문장 생성에 적합X

GPT-1대비 GPT-2 개선점

- GPT-1와 같이 트랜스포머에서 디코더 사용은 같음
- 보다 넓고 깊게 Multi-Head Attention 사용
- 기존 GPT-1 보다 학습된 **파라미터 수**와 데이터 양이 10배 이상

파라미터 최적화

Parameter Optimization

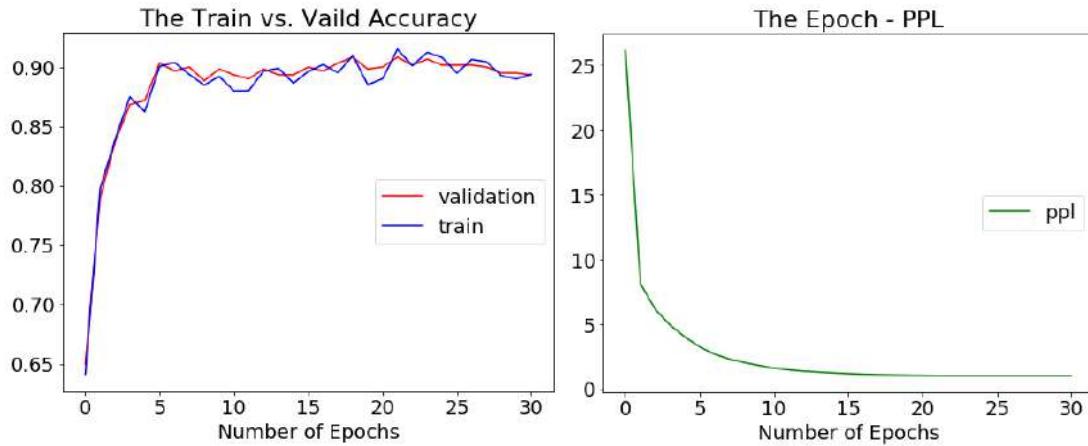
Argument	Default value	Modified Value	Description
Model	Open AI GPT	GPT2	Open AI에서 개발한 GPT, GPT2 모델
Num_candidates	2	6	Next Utterance의 후보군 선택지 수
Max_history	4	2	최대 History 수
Gradient_accumulation_steps	8	4	Optimize를 진행 시 GPU의 메모리 문제를 해결하기 위해 사용
Epochs	1	30	Epoch 수
Train_batch_size	4	2	1회 Epoch 시 Batch 크기(훈련 시)
Valid_batch_size	4	2	1회 Epoch 시 Batch 크기(검증 시)

출처: Thomas Wolf. "TransferTransfo: A Transfer Learning Approach for Neural Network Based Conversational Agents". 3page,
Zhang, S. "Personalizing Dialogue Agents: I have a dog, do you have pets too?". 2page

모델 선택

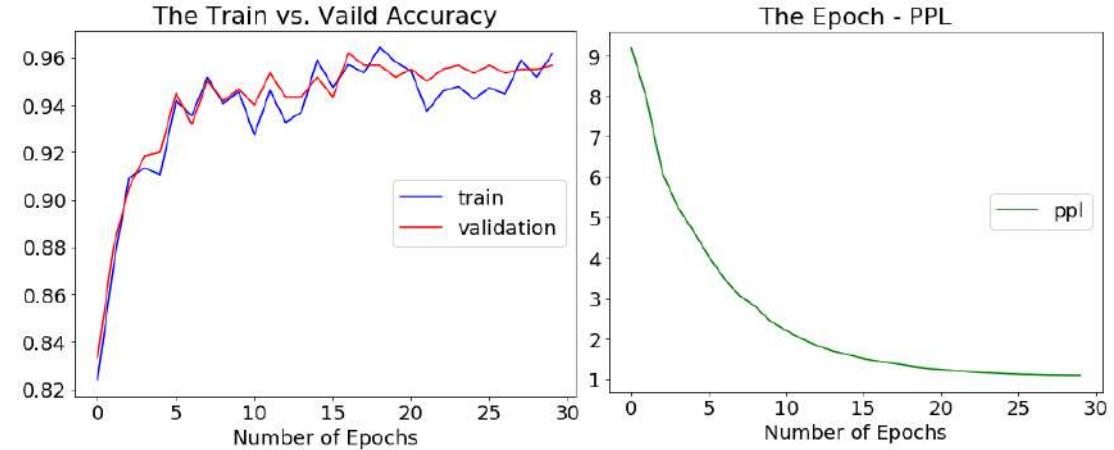
Model Selection

GPT-1



- 최적 Epoch : 2번째 Epoch 선택
- Hits@1 : 0.79
- Average PPL : 8.14
- 총 4시간 학습 (Ubuntu 18.04 LTS, GTX 2080Ti)

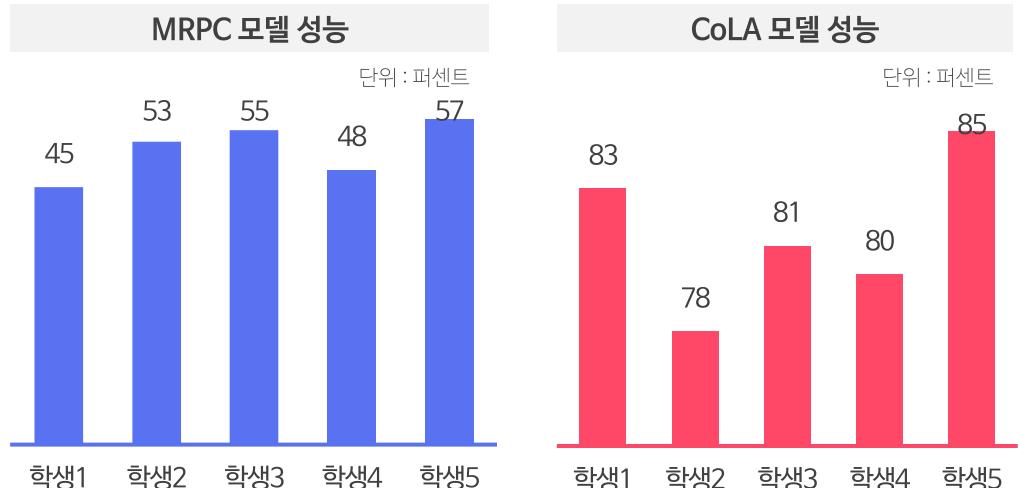
GPT-2



- 최적 Epoch : 2번째 Epoch 선택
- Hits@1 : 0.88
- Average PPL : 7.92
- 총 6시간 학습 (Ubuntu 18.04 LTS, GTX 2080Ti)

AFL 모델 테스트

AFL Model Testing



〈임의로 학생 5명을 선택하여 중간발표시 제작된 MRPC, CoLA 모델을 테스트〉

MRPC 사용 결과 : 평균 51.6%의 문장 유사도

CoLA 사용 결과 : 평균 81.4%의 문법 유사도

→ MRPC: 60%, CoLA: 80% 이상 대화 시 다른 Persona(상황)로
교체

디코더 설정

Decoder Setting

항목	Low	Medium	High
Top-k	180	70	0
Top-p	0.1	0.5	0.9
Temperature	1.9	1.2	0.6

〈Hugging Face에서 제공하는 디코더 세팅〉

- Decoder: Encoding 후, 출력 값을 다음 단어를 예측해가며 문장을 생성
- Top-k/Top-p : 기존의 Greedy/beam-search decoding 방식을 대체
→ 상위 k개의 토큰을 샘플링, 가장 인간이 사용하는 단어 분포와 유사
- Temperature : Softmax 함수에 Temperature Scaling 적용
- 여러 번의 실제 테스트 후 “Medium” 디코더 설정 사용

출처: *Hierarchical Neural Story Generation* by Angela Fan, Mike Lewis, Yann Dauphin
Language Models are Unsupervised Multitask Learners by Alec Radford, Jeff Wu, Rewon Child,
David Luan, Dario Amodei, and Ilya Sutskever

앱 개발

Application Dev.

REST API

- 대화형 Ai + MRPC 결합 후
- Python 프레임워크 Flask 사용하여 구현
- IDE : VS CODE



Flutter

구글에서 개발한 Android, iOS 빌드가 동시에 가능한 하이브리드 언어
→ Google Playstore와 App Store 배포 예정

각 모델 훈련 시간

Training Time

줌마영어 + Ai Hub 상황 데이터

GPU : GTX 2080TI
서버환경: Ubuntu 18.04 LTS (SSH사용)
소요시간: 총 6시간



CoLA

GPU : Tesla V100
서버환경: Google Colab
소요시간: 총 3시간



MRPC

GPU : Tesla V100
서버환경: Google Colab
소요시간: 총 3시간



04. Demo

프로젝트 구현

어플리케이션 시연

Demo Video

Hugging Face

비교

Comparison

기업 피드백

Enterprise Feedback



챗봇이 말을 생각보다 잘한다

전화영어 수업과 활용 시 정말 좋을 것

다양한 상황이 추가 되었으면 좋겠다

Ai 기능이 좀 더 정확했으면 좋겠다

향후 과제

줌마영어 회원 DB 제작 및 앱 활용

- 줌마영어 회원들의 정보를 입력할 수 있는 로그인 기능 및 DB 기능 추가
- 줌마영어 회원이 컨텐츠 학습(MRPC)과 문법 학습 (CoLA)의 데이터를 DB에 저장
- 이후, DB를 활용한 교육적 기능 추가적으로 개발

대화형 AI 학습량 증대

- 중간 발표 이후 총 3권의 줌마 영어 교재가 새로 출간됨
- 추가적으로 대응되는 챕터-상황 수를 늘려 데이터 수 증대
- 이후, 더 큰 데이터셋을 사용하여 새로운 학습 시행 예정

프로젝트 활용 방안

협업 기업의 수강생 활용

- 추가적인 모델 학습 후, 사용자의 피드백 반영
- 줌마영어 회원들이 쉽게 사용할 수 있도록 앱을 Google Play & App store에 등록하여 사용할 수 등록 예정

공모전 출전 및 스타트업 지원

- 현재 진행 중인 창업동아리를 활용하여
- 도전! K-스타트업 2020 AI리그
- NEXT RISE 스타트업 창업대회 출마 예정

업무 분담

김민우

- AFL 구축
- 앱 개발
- Database 구축
- REST API 구축
- 인터뷰 및 시장조사
- Presentation 관리
- 사전 NLP 논문 요약
- 제안서 발표

이승준(PM)

- Conversational AI구축
- Dialog Data Retrieval
- Missing Value Treatment
- NLP Paper 요약
- Text Pre-process
- 서버PC 관리
- 프로젝트 관리
- 최종 발표

김성진

- 아이디어 구체화
- 인터뷰 및 시장 조사
- 교육분야 논문 조사 및 분석
- 데이터 수집
- Flutter 프레임워크 담당
- 앱 개발
- 창업 동아리 일정 담당
- 중간 발표

감사합니다

04. Appendix

별 첨

참고 문헌

번호	이름
1)	Microlearning: A New Pedagogical Challenge, Theo Hug (2005)
2)	평생교육환경에서 인지적흥미와 학습자역할몰입의 조절효과: 자기결정성요인과 학습지속의지 간 관계를 중심으로, 임규연 (2016)
3)	이러닝 환경에서 몰입학습 증진을 위한 대화 기반 피드백 시스템의 개발, 정상목 (2007)
4)	이러닝 환경에서 몰입에 영향을 미치는 요인 연구 –상호작용 요인과 어포던스 요인을 중심으로, 이소영 (2019)
5)	학습자의 교육용 어플리케이션 활용 만족요인과 구매의도에 영향을 미치는 요인 연구, 장은지 (2012)
6)	Attention Is All You Need(2017)
7)	Deep contextualized word representations(2018)
8)	BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding(2019)
9)	Towards a Human-like Open-Domain Chatbot, D Adiwardana (2020)
10)	How to build a State-of-the-Art Conversational AI with Transfer Learning, Medium(2020)
11)	Hierarchical Neural Story Generation by Angela Fan, Mike Lewis, Yann Dauphin
12)	Deep Chit-Chat: Deep Learning for ChatBots, Wei Wu, Rui Yan(2019)
13)	Personalizing Dialogue Agents: I have a dog, do you have pets too?, ACL2018
14)	Training Millions of Personalized Dialogue Agents, EMNLP2018 Facebook
15)	Improving Language Understanding by Generative Pre-Training, Alec Radford(2018)
16)	Storage and retrieval processes in long-term memory., Shiffrin, R. M., & Atkinson, R. C. (1969).
17)	Language Models are Unsupervised Multitask Learners by Alec Radford, Jeff Wu, Rewon Child, David Luan, Dario Amodei, and Ilya Sutskever
18)	영어교과 보조도구로서의 AI 챗봇 분석 연구, 김혜영 외 3인 (2019)

선행 연구 요약(NLP 관련 논문)

Attention Is All You Need(2017)

- ✓ Attention-mechanism으로만 모델링
- ✓ 이전까지 존재하는 RNN기반 모델들의 Long term dependency problem 지적, 기존에 RNN시
- ✓ transformer에서는 학습 시 encoder에서는 각각의 position에 대해, 즉 각각의 단어에 대해 attention을 해주기만 하고, decoder에서는 masking 기법을 이용해 병렬 처리가 가능하게 됩니다

Deep contextualized word representations(2018.02)

- ✓ 문맥을 고려한 임베딩,
- ✓ biLM으로 부터 pre-trained 되어 만들어진 벡터를 ELMo 단어 임베딩에 사용한다 (biLM 의 벡터 각각 가중치들을 합하여 사용)
- ✓ 구축된 ELMo 모델에 task RNN에 전달하여 지도 모델 추가 가능

Improving Language Understanding by Generative Pre-Training

- ✓ embedding에 대한 이러한 문제점들을 Word-level 이 아닌 context-level로 구성
- ✓ Decoder에서 transformer 모델 사용
- ✓ 학습 시 Left-to-right 방향으로 진행
- ✓ Unsupervised Learning을 통한 Pretrained-LM의 구축
- ✓ Supervised Learning을 통한 Fine-tuning을 거쳐 Task에 맞는 NLP모델을 구성하는 가이드라인을 제시

BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding

- ✓ 단방향 Transformer를 사용했던 GPT-1과는 달리 Bidirectional Transformer를 사용해서 Pre-training을 더욱 효과적으로 수행
- ✓ MLM과 next sentence prediction이라는 2가지 방식으로 pre-training함
- ✓ 학습 시 GPT-1처럼 단방향이 아니라 양방향 위의 pre-training 방식 덕분
- ✓ Encoder 부분이 transformer 모델을 사용
- ✓ GPT-1 보다 좀 더 간단하게 fine-tuning 가능 레이어를 1개만 추가 해도 됨

선행 연구 요약(프로젝트 배경 관련 논문)

평생교육환경에서 인지적 흥미와 학습자역할몰입의 조절효과: 자기결정성 요인과 학습지속의지 간 관계를 중심으로

- ✓ 평생교육환경에서 학습자의 자기결정성 요인과 학습지속의지 간 관계에 인지적 흥미와 학습자역할몰입이 미치는 영향분석
- ✓ 첫째, 인지적 흥미는 자기결정성 요인 중 유능감과 학습지속의지 간 관계를 조절하는 것으로 나타남
- ✓ 둘째, 학습자역할몰입은 자기결정성 요인 중 유능감, 관계성과 학습지속의지 간 관계를 조절하는 것으로 나타났으며, 이 또한 대립적인 상호작용효과가 있는 것으로 나타남
- ✓ 이는 평생교육 학습자가 인지적 흥미를 느낄 수 있는 학습 경험을 제공하는 동시에 학습자로서 최선을 다할 수 있는 학습 환경을 구축해야 함을 시사

이러닝 환경에서 몰입학습 증진을 위한 대화 기반 피드백 시스템의 개발

- ✓ 교실수업에서 지속적인 면대면 피드백을 통해 학습동기를 부여하고 학습 성취 욕구를 충족시켜주어 몰입학습을 유도할 수 있으나 이러닝에서는 교실수업에 준하는 상호작용 수준에 미치지 못하고 있음
- ✓ 이러닝 환경의 문제점을 개선하는 방법으로 대화 피드백 시스템을 제안
- ✓ 몰입학습을 증진시키는 대화 기반 피드백을 적용
- ✓ 학습자에게 적극적인 학습참여 자세를 배양하고 학습의 흥미와 관심을 유지시켜 몰입학습을 유도하였음을 시사

선행 연구 요약(프로젝트 배경 관련 논문)

이러닝 환경에서 몰입에 영향을 미치는 요인 연구 -상호작용 요인과 어포던스 요인을 중심으로

- ✓ 몰입 이론을 기반으로 이러닝 이용자들의 학습 몰입에 영향을 주는 선행요인과 결과변수에 대한 관계를 분석
- ✓ 디지털 환경에서 학습자가 스스로 몰입되어 학습할 수 있는 선행요인을 살펴보고 점점 확대되는 이러닝 학습 콘텐츠 시장의 미래를 이해하고 기여할 수 있는 시사점 제시
- ✓ 학습자 통제감과 어포던스 요인 유희성이 유의한 결과가 나왔으며 몰입은 만족도에 유의한 영향을 주었다.
- ✓ 컴퓨터와 사용자 간의 상호작용에서 몰입에 영향을 미치는 요인들을 연구함으로써 이러닝 분야 외에 향후 다양한 분야에서 사람과 기계의 상호작용시 몰입을 통한 만족도 향상의 개발 방향을 제시

학습자의 교육용 어플리케이션 활용 만족요인과 구매의도에 영향을 미치는 요인 연구

- ✓ 스마트 기기의 대중화 및 교육용 어플리케이션에 대한 소비자의 수요 증가
- ✓ 교육용 어플리케이션 만족요인과 구매의도에 대한 사용자들의 요구를 분석
- ✓ 교육용 어플리케이션 보편화 시대를 대비해 개발 및 활용에 가이드라인 제시
- ✓ 적절성, 유희성, 편재성이 교육용 어플리케이션 구매의도에 긍정적인 영향을 미침
- ✓ 스마트 기기를 적극적으로 활용하여 참여협력 등 이른바 집단지성 구현을 촉진시키는 기능 개발이 향후 교육용 어플리케이션 제작 반영될 필요성 제시

시장 분석

넓은 범위의 시장 조사

교육분야 관심 키워드는 역시나 제 4차 산업혁명이다. 특히 인공지능 즉, AI 기술의 급속한 발달은 우리의 일상을 변화시키고 있으며, 산업은 물론 교육 분야에서도 AI 기술이 자리하기 시작하였다. 언제 어디서나 빠르고 정확하게 제공하는 AI 맞춤형 정보는 암기 위주의 지식학습 패러다임을 위협함과 동시에 외국어 교육 분야에서 독해나 문법학습 등 언어지식 전달 및 이해 위주의 기존 수업과 평가 방식이 아닌 회화 중심의 수업 방식으로 기존의 패러다임을 변화시키고 있다.

제 4차 산업혁명, 그 중에서도 인공지능과 모바일 앱 시장의 성장세를 파악하고 이들의 성장세가 교육 분야에 불러일으키고 있는 새로운 패러다임을 분석하고자 한다. 또한 AI 기술 수준에서 영어학습에 가장 활용가치가 높다고 판단되는 AI chatbot에 대해서 알아보고자 한다. 이를 파악하기 위해 AI chatbot이 왜 영어 학습에서 대두되고 있는지 조사하였다. AI chatbot이 대두되고 있는 근거에 대해서 살펴보고 현재 영어 교육시장에서 상용화되고 있는 AI chatbot사례들을 파악하는 한편, 각 기업마다 어떻게 차별화 시키고 있는지 알아보고자 한다.

경쟁업체 시장 조사 -트렌트를 바탕으로

최근 어학교육의 트렌드는 '성인 대상 초보 영어 시장 강세'와 'IT 기기와 연동한 스마트러닝'으로 요약된다. 본 조사에서는 현재 성인 대상 영어 회화 교육 기업과 IT 기술 중 AI를 이용한 영어 회화 교육 기업에 대해서 조사하였다. 국내 성인 대상 교육 시장의 규모는 약 2조 5천억 원이며, 그 중 외국어 교육은 약 1조 8천억 원으로 전체 성인교육 시장에서 압도적인 우위에 있다.

이 중 성인 대상 영어 교육 시장 규모는 7,200억 원을 차지한다. 최근 생활 패턴과 인식의 변화로 영화감상, 여행 등 실전 영어의 필요성에 의한 성인 대상 초보 영어 회화 시장이 강세를 보이고 있다. 성인들은 직장과 육아 등으로 인해 오프라인 수강보다 시간에 구애 받지 않는 온라인 수강을 선호한다. 이렇듯 시간 효율성에 대한 소비자들의 니즈가 높아져 온라인/모바일 기반의 영어 회화 교육이 증가하는 추세이다. 새로운 니즈의 발생으로 토익, 독해, 문법 강의 등의 전통적인 오프라인 영어 시장 이외에 새롭게 온라인 회화 영어 교육 시장이 등장했다. 현재 오프라인 영어 교육 시장에서 온라인 영어 교육 시장으로 패러다임이 변화하면서 온라인/모바일 시장을 선도하는 기업들이 나타났다. 대표적으로 시원스쿨, 야나두, 뇌새김, 스피킹맥스 등이 30,40세대의 영어 말하기에 대한 갈망을 제대로 읽고 새로운 패권자로 나타났다.

Hits@1

맞은 확률 = 정확도

		Actual	
		Positive	Negative
Predicted	Positive	True Positive	False Positive
	Negative	False Negative	True Negative

- Accuracy = $(TP+TN)/(TP+FP+FN+TN)$

Perplexity

문장의 길이를 반영하여 확률값을 정규화한 값

$$\begin{aligned} PPL(w_1, w_2, \dots, w_n) &= P(w_1, w_2, \dots, w_n)^{-\frac{1}{n}} \\ &= \sqrt[n]{\frac{1}{P(w_1, w_2, \dots, w_n)}} \end{aligned}$$

- 문장의 확률값이 분모에 들어가 있으므로, 확률값이 높을수록 PPL은 작아집니다.
- PPL은 수치가 낮을수록 좋다.
- 보통 n-gram의 n이 클수록 더 낮은 PPL을 보여줌

프로젝트 방법론

목적별 사용 Framework/ Language		
Framework	Language	목적
Jupyter note book	Python	64GB
Flutter	Dart	300g
Android studio	5.8형	5.8형
Git/Github	4000화소	4000화소
Docker	30frame	30frame

목적별 사용 Library	
Library	목적
Numpy	Data Analysis
Pandas	Data Analysis, Data Preprocessing
Matplotlib	Data Analysis
Requests	Making HTTP request
Urllib request	Opening URLs
Torch	Deep Neural Network
Tensorflow	Deep Neural Network