

데이터 분석 캡스톤 디자인

문제 정의

기존 서비스

1. [애플워치 카톡 답장 방식]
2. [애플워치 메세지 답장 방식]
3. [갤럭시 워치 메세지 답장 방식]
4. [카톡 말투 교정 AI]
5. 애플워치 답장 리스트
 - [메세지 앱]
 - [카카오톡]
6. 갤럭시 워치 답장 리스트
 - [메세지 앱]
 - [카카오톡]
7. 갤럭시 자동 응답

기존 연구

1. 답장과 관련된 사회적 특성
2. 답장 제안 관련 연구
3. 스마트워치에서의 텍스트 입력

예상 시나리오 (주변인 의견)

TODO

예상 데모

문제 정의

스마트 워치는 휴대폰을 지니고 있지 않더라도 전화, 메세지 등의 연락에 즉시 확인할 수 있다는 장점을 가지고 있다. 하지만 타이핑을 하기에는 인터페이스가 매우 작기 때문에 메세지에는 쉽게 답장하기 어렵다.

이를 해결하기 위해 애플워치와 갤럭시 워치에서는 ["좋아", "알겠습니다", "응"]이라는 답장 리스트, 쿼티 키보드, 음성 받아쓰기, 손글씨 입력 등의 기능을 지원하고 있고, 카톡에서는 음성녹음을 전송하거나, 이모티콘을 사용하여 답장하도록 제안하고 있다.

하지만 음성 녹음을 통한 답장은 주변이 시끄럽거나, 사람이 많은 경우 품질 및 사회적 시선으로 인해 사용하기 불편할 수 있고, 답장 대상에 따라 사용할 수 없을 수 있다. (예를 들어 직장 상사)

또한 제안되는 답장 리스트는 택배기사, 가족, 여자친구 모두 똑같은 답장 리스트를 제안하고 있으며, 그 내용 또한 단순하기 때문에 상대방의 연락에 대한 나의 대답을 정확히 표현할 수 없고 상대방과 나의 관계에 따른 답장 방식 차이를 녹여낼 수 없다.

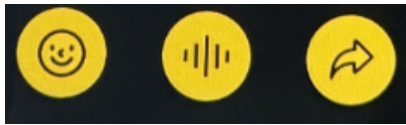


핵심 문제 상황

1. 스마트 워치의 작은 인터페이스 → 답장 타이핑 불가능
 2. 현재 사용되는 답장 방식 : 음성녹음, 이모티콘, 답장 리스트 제안
 3. 음성 녹음 : "주변 상황, 상대방에 따라 사용하기 어려움"
 4. 답장 리스트 제안 : "상대방이 누구던지간에 다 똑같은 단순한 답장만 가능"
- ⇒ 현재 사용되는 방식으로는
"내가 전달하고 싶은 답장 내용을 상대방과의 관계에 맞춰서 보낼 수 없다!"

기존 서비스

1. [애플워치 카톡 답장 방식]



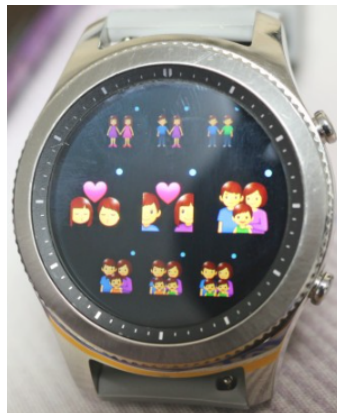
2. [애플워치 메시지 답장 방식]

Apple Watch에서 메시지에 답장하기



- 손글씨 입력, 쿼티 키보드, 텍스트 받아쓰기, 이모티콘, iPhone으로 텍스트 입력

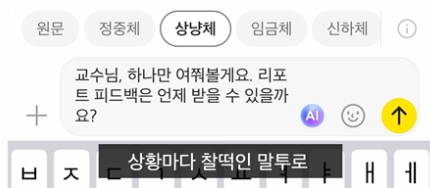
3. [갤럭시 워치 메시지 답장 방식]





4. [카톡 말투 교정 AI]

AI로 고민되는 카톡 말투도 쉽게, 말투 변경하기



⇒ 내 말투로 변경시키는 서비스는 없을 뿐더러, 스마트워치에의 답장 제안 서비스에는 도입된 적이 없다.

5. 애플워치 답장 리스트

[메세지 앱]

좋아

알겠습니다

응

예

아니

아니요

고마워

감사합니다

나중에 전화해도 될까?

나중에 전화드려도 될까요?

[응답 추가] → 딱히 개수 제한이 없는 것 같음

[카카오톡]

네

아니요

감사합니다

어디야?

나중에 연락할게

회의 중입니다

운전 중입니다

지금 가는 중입니다

다른 용무 중입니다

지금은 답장을 할 수 없습니다.

[위의 default 메시지를 자신의 응답으로 커스터마이징 가능]

6. 갤럭시 위치 답장 리스트

[메세지 앱]

▼ “택배 알림에 대한 답장 리스트”

네 알겠습니다

아 그래?

네. 그렇게 해주세요

좋아

아니

괜찮습니다

싫어

고마워

매우 감사합니다

가는 중이야.

무슨일이예요?

나중에

멋져

그러게

오오!

정말?

이런!

[카카오톡]

▼ “아무말에 대한 답장 리스트”

헐

아 그래?

응

네

네, 그렇게 해주세요

좋아

아니

괜찮습니다

싫어

고마워

매우 감사합니다

가는 중이야

무슨 일이에요?

나중에

멋져

그러게

오오!

정말?

이런!

▼ “안녕하세요에 대한 답장 리스트”

네 안녕하세요

오

네

네, 그렇게 해주세요

좋아

아니

괜찮습니다

싫어

고마워

매우 감사합니다

가는 중이야

무슨 일이에요?

나중에

멋져

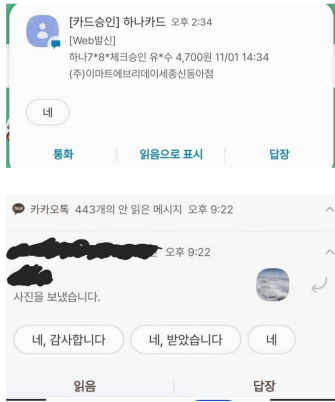
그러게

오오!

정말?

이런!

7. 갤럭시 자동 응답



기존 연구

1. 답장과 관련된 사회적 특성

"I Can't Reply with That": Characterizing Problematic Email Reply Suggestions [CHI 21']

: 사용자 인터뷰를 통해 이메일 답장 제안에서 문제가 될 수 있는 부분들 (사회적 맥락, 구조적 특징)을 식별하고 영향을 정량화한 연구

▼ 논문 내용

문제 상황

: 텍스트 예측은 글쓰기 부담을 줄여주지만, 품질을 우선시하다보니 문법, 철자, 문장 구성은 고려하지만, 문맥을 고려하지 않아 무례하거나 부적절한 제안을 포함시키게 된다. 또한 언어의 이론적 틀은 의사소통의 사회적 의미를 반영하지 못한다.

실험 1: 이메일 답장 제안이 문제가 되는 주제 식별

1) 사용성

- 유용성: "간단한 확인이 필요한 이메일에 유용함" (추가적인 설명 불필요할때)
- 사용자 경험: "응답을 선택하면 바로 전송될 것 이라는 오해가 있었음"
- 주체성: "내가 답장을 업데이트하고 결정하고 싶다"

2) 이메일 주제 (구조적 특징, 진정성, 의미적 특성)

- 구조적 특징: 메일을 시작하는 말, 끝내는말 → 더 공식적, 전문적으로 만들
- 진정성: "너무 형식적이면 감정적 측면이 사라진다"
- 의미적 특성
 - 혼란스러움 (메일의 질문과 일치하지 않는 답장일 때, 시제가 맞지 않을때)
 - 무례함 (짧거나 간결한 회신)
 - 불일치 (긍정적인 메일에 슬프게 답장)
 - 지나치게 긍정적인 표현을 싫어함 (사회언어학적으로 긍정적인 단어가 많기 때문에 이메일 제안도 긍정적 제안이 많지만, 모든 상황에 긍정적 회신이 필요한게 아님)

3) 사회적 맥락

- 관계 유형: 사회적 계급에 따라 달라져야한다 ex) 상대에 따라 여는말, 닫는말을 포함시킴
↔ 사회적 관계더라도, 2~3년 함께한 팀원들에게는 여는말과 닫는말을 생략하기도함
- 문화적 요인: 문화에 따라 선호하는 글쓰기 스타일 및 구조적 특징이 있다.

실험 2 : 답장에서의 위의 요인(구조적 특징, 사회적 관계)들의 영향 식별

: 다양한 시나리오를 세팅 → 제안된 답장을 선택, 평가, 수정하도록 함

1) 문화적, 긍정적, 성을 가정한 답변은 일반적으로 평가가 높음 + **답변을 그대로 보내는 경향이 있다**

2) 비전문적, 혼동스러운, 부조화, 무례한 답변은 일반적으로 낮음 + **답변을 그대로 안보냄**

“이때 한 척도가 부정이면, 다른 척도들은 모두 긍정이었음”

3) 비전문적인 시나리오가 그대로 보내지는 경우는 상사, 동료, 멘토보다 친구, 부모, 형제일 때 5~11% 적었음

4) 상사에게 보내는 시나리오 일때 답변이 가장 적절하지 않다고 평가됨 + **답변을 그대로 보내는 비율이 멘토와 함께 가장 낮았음 (35.6%, 36.4%)**

5) 형제, 자매에게 보내는 답변은 그대로 보내는 비율이 가장 높았음 (39.2%, 39.6%)

6) 구조적 특징 (시작하는말, 끝내는말 유무)는 딱히 영향 X

7) 무례한 메일일 경우 가장 많이 (평균 6.9단어) 수정됨, 빈도도 높다



논문 요약

: 답장을 제안할 때 맥락의 올바름과 사회적 관계를 고려해야 한다.

인용할만한 내용

*

문화적, 긍정적, 성을 가정한 답변은 평가가 높고, 그대로 보내는 경향이 있다

⇒

상대방에 대한 정보에 맞춰서 답변을 제안하면 성공률이 높을 것

* 사회적 관계 (상사, 멘토, 동료)에 있는 상대방에게는 답장의 공손함이 중요함

“하지만 사회적 관계 안에서도 인사말을 생략하는 등의 차이가 있음”

⇒

상대방 별 답장을 사회적 관계, 그 안의 개인적 관계까지 확장시켜야 한다.

* 형제, 자매 등의 가까운 지인에게는 속어를 쓴다 + 너무 형식적이면 감정적 측면이 사라지고, 짧으면 불쾌하다.

⇒ 형제 자매 등의 편한 관계에서 감정적 상호작용을 위해 속어 등의 그들 간의 커뮤니케이션을 답장에 적용해야하고, 너무 짧아서는 안된다.

The Study of AI-Mediated Communication and Socio-Cultural Language-Related Variables: Gmail Reply Suggestions [Applied Artificial Intelligence '23]

▼ 논문 내용

문제 상황

: ERS (Email Reply Suggestion) 서비스는 상황, 대인관계, 주제성, 긍정성, 성별, 언어 및 종교 등과 관련하여 몇 가지 문제를 야기하고 있다.

→ 기존에는 ERS 서비스에 대한 평가를 정량적으로만 측정하였지만 사회/언어적인 변수를 고려 할 필요가 있음

연구 방법

1) Gmail 사용자들에게 자체 제작한 설문을 배포한다.

2) 직업, 전문분야, 연령, 성별로 집단을 나누어 분석한다.

연구 결과

(표 및 그래프 추가 예정)

“ERS 서비스에 대한 인식과 사용성은 사용자의 사회 집단에 상관관계가 있다”



논문 요약

: 이메일 답변 제안 시스템을 사회/언어적 변수와 관련하여 더 잘 구현할 것을 권장하고 관련 연구를 제안함

인용할만한 내용

*

ERS (Email reply suggestion) 시스템에 대한 인식과 사용성은 여러 변수로 구분한 사회/언어적 집단에 따라 달라 짐
* 사용자의 사회적 위치를 고려해야한다.

⇒

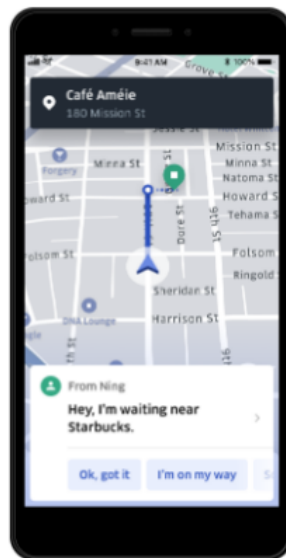
우리가 제작 할 답변 제안 시스템도 사용자와 상대가 소속된 사회 집단에 영향을 받도록 만들면 좋을 것이다

⇒ 답변 제안 시스템을 나와 상대방의 관계 및 소속된 사회 집단까지 고려하여 작동되게 하는 것이 과연 구현할 수 있고 효율적인지 생각 해 볼 필요가 있을 듯 함

2. 답장 제안 관련 연구

OCC: A Smart Reply System for Efficient In-App Communications [KDD '19]

: UBER 앱에서 탑승자와 운전자사이에 **원클릭 채팅 시스템**을 구현한 논문



▼ 논문 내용

문제 상황

: 우버에서는 운전자와 승객을 연결시키기 위해 채팅 기능을 구현했으나, **운전자가 운전중에 메시지를 입력해야한다는 문제**가 있었음

→ 대화 상황에 따라 **동적으로 제안되는 스마트 답장**을 통해 커뮤니케이션할 수 있도록 원클릭 채팅 앱 구현

연구 방법

1) 인앱 메시지 데이터로 메시지 의도 학습 ex) “어디세요” → “현재 위치”라는 의도를 감지

2) 해당 의도와 관련성이 가장 높은 상위 4개의 답장 제안

ex) 예, 가고있는 중, 죄송합니다, 교통 체증 중, 픽업 주소에 있습니다, 전화 주세요

의도 학습 방법 : 빈도 기반 / cnn 기반 / 임베딩 + nearest neighbour classifier

연구 결과

의도 예측 모델 정확도

- 빈도 기반 : 15.5 %
- CNN : 77.2 %
- NNC : 75.9 %
- WORD-NNC : 75.6 %

"4개를 예측했을 때 적어도 하나 있을 정확도는 밑의 3개 모델 모두 90% 이상의 정확도를 보임"

실제 배포 → 영어권 국가에서의 여행 중 23%가 인앱 쌍방 통신을 사용하고, 그중 71%이상이 OCC를 활용함



논문 요약

:

응답을 직접 예측하는 것이 아니라, **의도 감지와 응답 검색으로 프로세스를 나누어서** 응답 제안을 구현함

인용할만한 내용

-

응답을 직접 예측하는 것이 아니라, 메시지에서 의도를 감지한 후 의도에 맞는 답장을 매핑시켜 제안했고, 실제 상용화 까지 성공함

→ 우리 연구에서도 비슷한 맥락으로 채팅으로부터 의도(맥락)를 추출하고, 추출된 의도와 카톡방의 top-n 답장 + 변형 답장을 매핑시킴으로써 상용화가 가능한 답장을 제안할 수 있다.

* 물론 기존 연구에서는 탑승자와 운전자 사이의 간단한 컨텍스트만 다뤄서 가능했던 것일 수 있다.
(우리꺼가 더 어려울 것)

LLM-based Smart Reply (LSR): Enhancing Collaborative Performance with ChatGPT-mediated Smart Reply System ['23]

Smart Reply: Automated Response Suggestion for Email [KDD '16]

Efficient Natural Language Response Suggestion for Smart Reply [GOOGLE '17]

▼ 논문 내용

문제 상황

: 현재 일반적으로 사용되는 Seq2Seq 모델이 효율적인지 확실하지 않으며 학습이 느리고 복잡하다.

→ (Seq2Seq 모델에 주로 사용되는) LSTM 네트워크보다 빠른 feed-forward 네트워크와 비용이 적게 드는 n-gram 임베딩을 기반으로 구축된 네트워크를 사용한다.

연구 방법

- 1) 사용자가 이메일에 신속하게 응답할 수 있도록 짧은 응답 제안을 제공한다. (Via inbox by Gmail)
- 2) 모델을 발전시키며 온라인 및 오프라인으로 평가한다.

모델 발전 과정 : 내적 모델 - joint scoring 모델 - two pass setup - single pass setup

연구 결과

Conversion / Latency (relative to Seq2Seq)

- two pass setup : 104% / 10%
- single pass setup : 104% / 1%

“최종적으로 Seq2Seq 시스템보다 높은 품질과 적은 Latency 로 답장 제안을 구현하는데 성공함”



논문 요약

: 기존보다 효율적인 자연어 응답 제안을 위한 기계 학습 방법을 제시한다.

인용할만한 내용

*

기존에 사용하던 Seq2Seq 모델과 LSTM 네트워크는 학습이 느리고 복잡하다.

⇒

LSTM 네트워크보다 빠르고 효율적인 n-gram / feed-forward 네트워크를 사용

* 내적 모델 - joint scoring 모델 - 2 pass setup - single pass setup 의 순서로 모델을 발전

⇒

원래의 Seq2Seq 시스템보다 품질이 향상되고 비용과 Latency 가 줄어든 시스템을 만들 수 있게 됨

* 구현단계에서 참고할 수 있을 듯 함

3. 스마트워치에서의 텍스트 입력

Text Input on a Smart Watch [IEEE Pervasive Computing '14]

Using a smart-watch as a input device for text [IEEE Pervasive '14]



논문 요약

:

이 논문에서는 스마트워치에서 텍스트 입력이 제한됨에 따라 상호 작용이 제한되는 점을 지적하고, 해결방안으로 소형 텍스트 입력방식을 제안한다.

인용할만한 내용

-

현재 스마트워치의 텍스트 입력을 위해 고안된 방안으로는 풀 쿼티 키보드가 스마트 시계 화면을 채우기 위해 축소되는 Zoomboard 방식이 있다. Zoomboard는 한 번 탭하여 키보드 영역을 확대하고 두 번째 탭하여 해당 영역에서 문자를 선택하는 방식이다.

→ 텍스트 입력은 많은 스마트폰 애플리케이션의 핵심 구성 요소이다. 스마트워치는 상당히 많은 관심을 받고 있지만 텍스트 입력이 없으면 상호 작용이 실망스러울 정도로 제한된다.

→ 게시글, 짧은 메시지 및 이메일을 볼 수 있지만 동일한 장치에서 답장을 보낼 수는 없다는 단점이 있다.



예상 시나리오 (주변인 의견)

1. 고정 스케줄 (강의, 회의, 알바)
2. 운동 (러닝, 자전거...)
3. 운전
4. 게임
5. 요리
6. 컴퓨터
7. 스마트폰이 멀리 떨어져있을 때 EX) 침대에 누워있는데 핸드폰은 책상에 있을 때

TODO

1. watch 로 답변을 하는 상황들은 어떤 상황인가

- a. 위 예시들 처럼 두 손이 부자유스럽거나 바쁜 상황일 것 (대화가 이어지면 안됨)
 - i. 현재 대화하기 힘든 상황이라고 사유를 알려 주는 경우가 대부분일 것인데 대답해야 하는 상대방에 따라 현재 내 사유는 변하지 않을 것이다. 그러므로 상대방에 따라 대화 내용이 바뀌는 것은 아니고 말투만 바뀌는 것이 맞을 것.
- b. 위 상황들을 어떻게 인지하게 할 것인가
 - i. 고정 스케줄은 에타 시간표를 통해 알 수 있고, 운전은 GPS이동 속도, 게임 중 또한 인지할 수 있음
 - ii. 운동중과 같은 스마트워치가 인식하는 것을 실시간으로 알아낼 수 있는지 추가 조사 필요.. (아직 못 찾음)

2. 데모가 어떻게 구현되어야 하는가

- a. 카톡 채팅방과 유사한 채팅방 시스템을 구현
 - 실험자에게 시나리오를 부여하고, 해당 시나리오안에서 상대방에 맞춰서 제안된 응답중 어떤 것을 선택할지 혹은 선택하지 않을지를 확인
- b. 어떤 모델 사용해서 관계분석 구현할 것인가

▼ 관련자료

▼ tensorflow lite smart reply - 답장제안

https://github.com/tensorflow/examples/tree/master/lite/examples/smart_reply/android/

▼ on-device smart reply - 왓치 등 기기에서의 답장제안

<https://blog.research.google/2017/02/on-device-machine-intelligence.html>

▼ GPT 파인튜닝 - 말투교정

<https://thinkingboy83.tistory.com/entry/GPT-35-파인튜닝-API로-자신만의-모델을-만들어보자>

<https://velog.io/@tmddn0311/chatbot-like-me-2>

▼ smilegate - 말투 선택

https://github.com/smilegate-ai/korean_smile_style_dataset

→ 구현 : 스마트워치 앱으로 채팅방 만들어서 답장을 제안하는 건 가능함

: GPT api를 이용하여 말투교정하는데 과금이 발생함

c. 플랫폼, 툴, 언어 선정

→ Android studio, java

d. 답변 추천을 Yes / No 편향되게 할 것인가

- a. 긍정답변 추천할 경우로 판단되었을 때 부정답변도 한두가지 추가해서 보여 줄 것인가

e. 실제 구현된 RS (Reply Suggestion) 시스템/서비스에서 많이 사용한 대답 등에 대한 통계가 있는지?

- a. 찾지 못함

예상 데모

