

ADA: 인공지능 챗봇 플랫폼

Lead Scientist

- John Park
 - UC. Berkeley EECS / Cognitive Science
 - Symantec / HP Principal Engineer
- Kaggle World Ranking 24th
- Expert @ Deep-learning

특징

- 페이스북 API 연결
- 머신러닝과 의도분석에 기반한 대화
- 채팅의 이관 가능성에 중점을 둔 알고리즘
- Siri와 유사
- 11월 초 8퍼센트에 적용
 - (15개 질문에 대응)



Chatbot framework : How to use (예시)

1. Admin이 Chat Thread 를 삽입
2. Intention Tag 를 삽입
3. 학습 및 새로운 Intention Suggestion
4. Template Engine 작성
5. DB와 연결
6. Chatbot API와 연결

➔ 고객은 자동화된 응답을 받게 됨

System UI : Main Panel

ADA Admin Panel v0.2

Core > Chat threads

결과 3543개 나옴 총 10427건

<input type="checkbox"/>	ID	First question	FQI	FQI Change	SFQI 1	SFQI 1 C	SFQI 2	SFQI 2 C
<input type="checkbox"/>	20259	보안상 삭제됨	14. 대출; 문의; 프로세스	-	14. 대출; 문의; 프로세스	4.4587382906457425	11. 대출; 문의; 대출필요한것같은	1.564701629848024
<input type="checkbox"/>	20258	보안상 삭제됨	41. 문의(인사)	-	45. 투자; 안심편드란?	4.832765306287287	03. 기타	1.3845905427546523
<input type="checkbox"/>	20257	보안상 삭제됨	14. 대출; 문의; 프로세스	-	17. 대출; 서류; 휴대폰; 법인명의	8.098639493390461	43. 사용;버그	2.9155352230906333
<input type="checkbox"/>	20256	보안상 삭제됨	24. 대출; 진행; 서류확인요망	-	24. 대출; 진행; 서류확인요망	6.18531183203294	14. 대출; 문의; 프로세스	3.767612895073667
<input type="checkbox"/>	20255	보안상 삭제됨	43. 사용;버그	-	43. 사용;버그	7.59108297976997	03. 기타	1.542406762762152
<input type="checkbox"/>	20254	보안상 삭제됨	03. 기타	-	03. 기타	1.8850541290937914	43. 사용;버그	1.7870493979342905
<input type="checkbox"/>	20253	보안상 삭제됨	41. 문의(인사)	-	43. 사용;버그	0.9842568003096074	03. 기타	0.8092930412828813
<input type="checkbox"/>	20252	보안상 삭제됨	03. 기타	-	03. 기타	4.885630183463672	22. 대출; 제출방식	1.8042458454654844
<input type="checkbox"/>	20251	보안상 삭제됨	-	-	11. 대출; 문의; 대출필요한것같은	3.180941990276436	41. 문의(인사)	2.6589353309717794
<input type="checkbox"/>	20250	보안상 삭제됨	22. 대출; 제출방식	-	22. 대출; 제출방식	7.418299331997671	14. 대출; 문의; 프로세스	3.59700455825272
<input type="checkbox"/>	20249	보안상 삭제됨	41. 문의(인사)	-	41. 문의(인사)	1.9842306156650005	03. 기타	1.3845905427546523
<input type="checkbox"/>	20248	보안상 삭제됨	14. 대출; 문의; 프로세스	-	25. 대출; 프리랜서신청방식	5.793493377440993	14. 대출; 문의; 프로세스	4.185550493018575
<input type="checkbox"/>	20247	보안상 삭제됨	14. 대출; 문의; 프로세스	-	14. 대출; 문의; 프로세스	6.451821823510787	11. 대출; 문의; 대출필요한것같은	5.34433674968943

100 중 아무것도 선택되지 않았습니다. 실행

System UI : Real-Time analyzer

Chat Thread Analyzer

Chat Thread Status : Running

Chat Log

Q: Hello
A: How are you?
Q: Do you need any help?
A: Yes please

Question

Sample Text

Proceed

Answer

1) 그렇고 그렇고 그렇죠
2) 앓 감사합니다.
3) 알겠습니다.
4) {{call}}
5) Sample Text

Proceed

Intention Status

31. 대출: 일반 0.45
53. 대출: 상환 0.51
21. 투자: 예치금 0.21
01. 계정: 암호 0.21
11. 일반: 버그 0.11

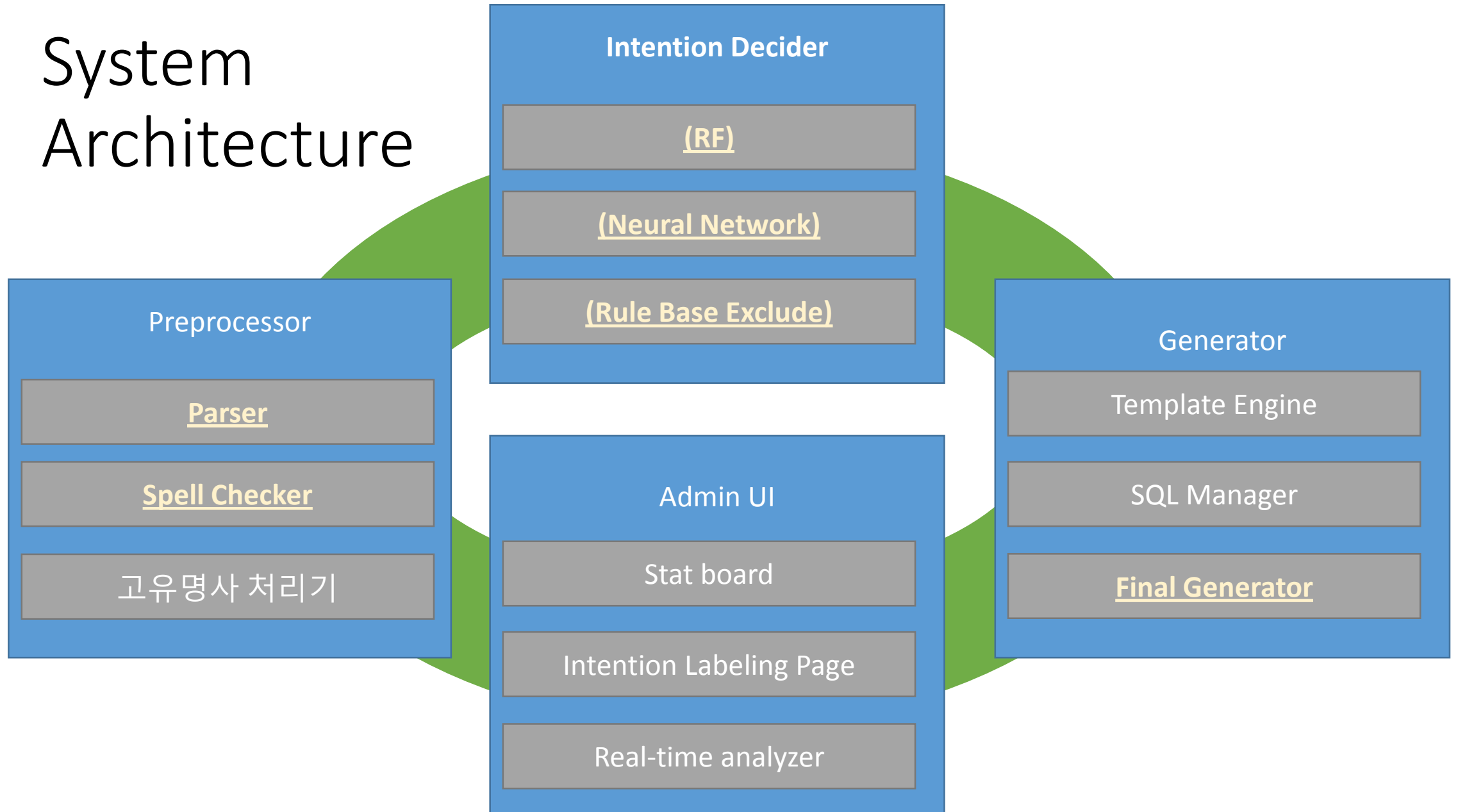
Intention

Nothing

Change

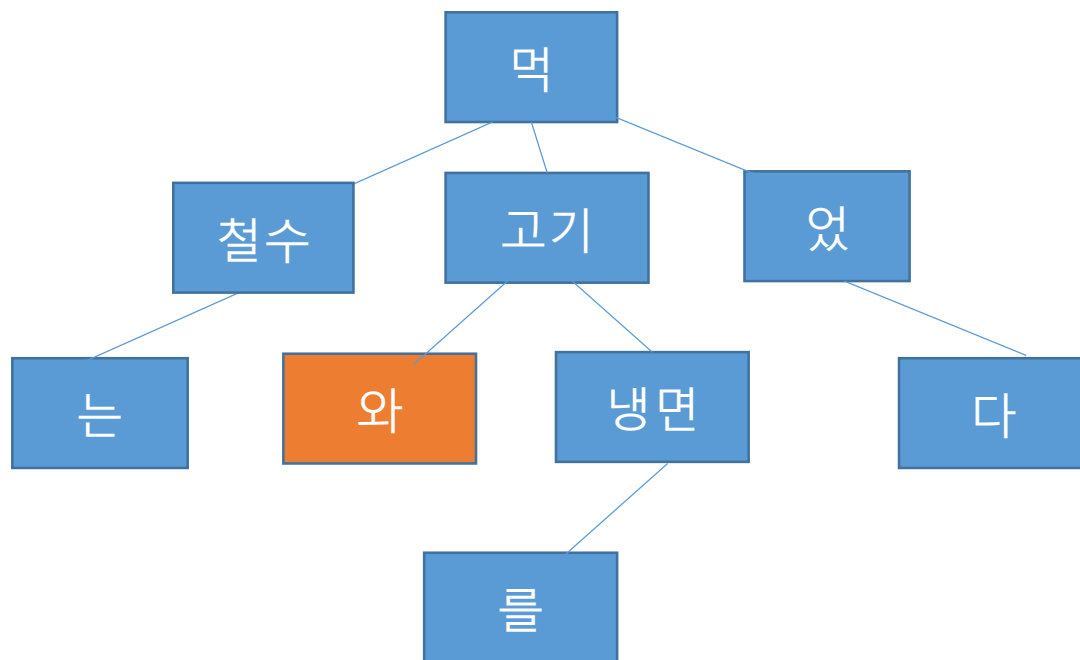
Finish

System Architecture



Parser (open api 형태로 공개)

사용 예 : 철수는 냉면과 고기를 먹었다



Parser

- 의도를 정확하게 캐치해내기
 - 시제가 매우 중요
 - 팩스로 서류를 제출하려합니다 → 네, 팩스번호는 02-000-0000 입니다.
 - 팩스로 서류를 제출했습니다. → 네, 확인결과 수신을 완료하였습니다.
- 예)
 - “전화번호와 팩스번호는 무엇인가요?”
=
“전화번호는 무엇인가요, 그리고 팩스번호는 무엇인가요”

→ 동일한 벡터를 가져야함

고유명사 처리기

대한민국 1등 숙박 어플
여기어때.



 **아놀리**

Chat example

- NLU
- NLP
- Parser
- Template Engine
- Context Awareness
 - Customer
 - Deal
 - Time

- Chat started on 2016-08-03 03:14 AM UTC
- (03:14:13 AM) Q: 안녕하세요. 호텔왕 펀딩에 관해 여쭙보려고요
- (03:14:24 AM) A: 네 고객님 어떤 점이 궁금하신가요?
- (03:14:24 AM) Q: 투자 진행은 언제 되는지요?
- (03:14:33 AM) A: 오늘 오후 1시 시작됩니다.
- (03:15:17 AM) Q: 투자 금액 전액이 펀딩이 안되면 없어지는 건지요?
아니면 투자액만으로 펀딩이 진행되는지요?
- (03:15:35 AM) A: 투자금액 전액 펀딩을 목표로 하고 있습니다.
- (04:00:28 AM) Q: 호텔왕 저는 대기로 되어 있는데 다른 분들은 펀딩이 되는지요
- (04:00:40 AM) A: 홈페이지 새로 고침 해 주셨나요?
- (04:00:53 AM) Q: 네 이제 되네요
- (04:00:59 AM) A: 네 감사합니다.

- Chat started on 2016-08-03 03:14 AM UTC
- (03:14:13 AM) Q: 안녕하세요. 호텔왕 펀딩에 관해 여쭙보려고요
- (03:14:24 AM) A: 네 고객님 어떤 점이 궁금하신가요?
컨텍스트 한정됨. 의도 불명확. 되물음.
- (03:14:24 AM) Q: 투자 진행은 언제 되는지요?
- (03:14:33 AM) A: 오늘 오후 1시 시작됩니다.
컨텍스트 내부 정보
- (03:15:17 AM) Q: 투자 금액 전액이 펀딩이 안되면 없어지는 건지요?
아니면 투자액만으로 펀딩이 진행되는지요?
- (03:15:35 AM) A: 투자금액 전액 펀딩을 목표로 하고 있습니다.
추가적 Intention. 지속대화. 컨텍스트 내부 정보
- (04:00:28 AM) Q: 호텔왕 저는 대기로 되어 있는데 다른 분들은 펀딩이 되는지요
- (04:00:40 AM) A: 홈페이지 새로 고침 해 주셨나요?
컨텍스트 변화 (시간). 대답 변화. 컨텍스트 내부 정보.
되물음
- (04:00:53 AM) Q: 네 이제 되네요
- (04:00:59 AM) A: 네 감사합니다.
인텐션 Off

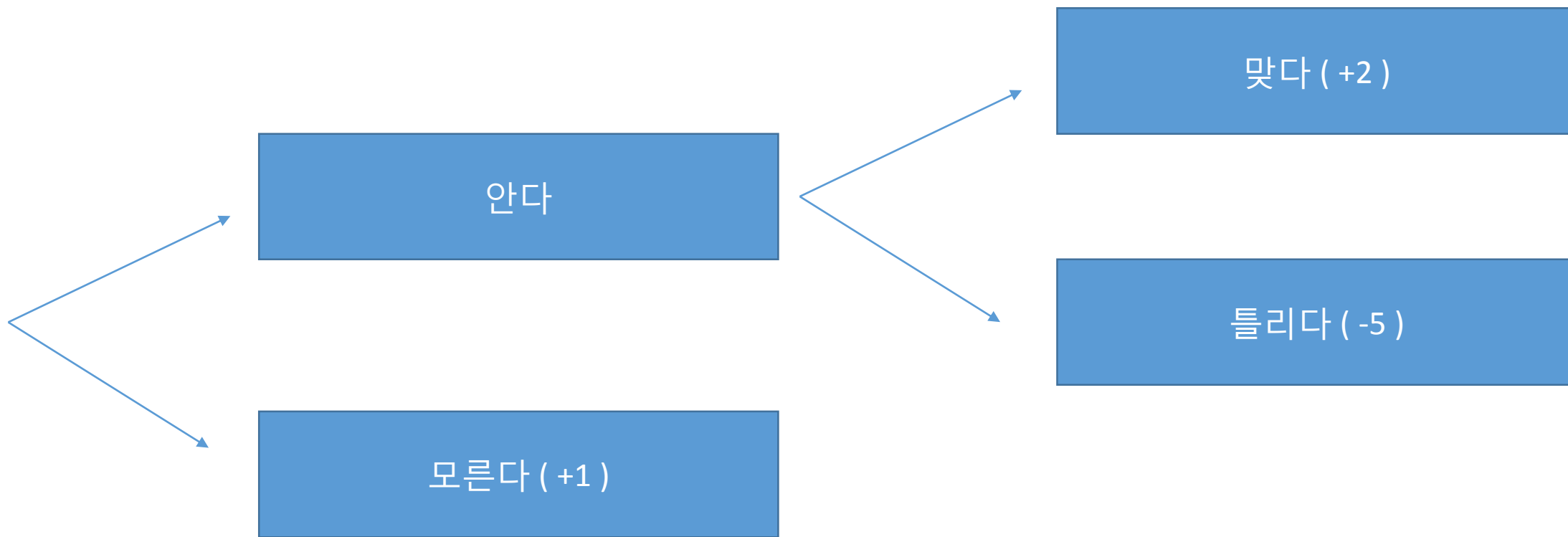
Real time machine learning in industry

- 맞을 확률 vs 틀릴 확률 vs 모를 확률

Q: 투자 수익률은 얼마인가요?

맞는 대답	8퍼센트의 투자수익률은 투자자마다 차이는 있지만 통계상 최종 세전 현재 약 9%입니다. 이는 부도, 연체를 포함한 수익률입니다.
틀린 대답	8퍼센트에는 약 1만명의 투자자가 투자를 하고 계십니다.
모르는 대답	네, 상담사를 연결해드리겠습니다.

➔ 맞는 대답을 하는 것보다 틀리지 않은 대답을 하는 것이 훨씬 중요!



AI model : Intention Decider

- NN
- CNN
- RNN
- RF
- Rule-based

Intention Hierarchy & Intention Decider

- Error 을 세게 주는데 이용
 - 올바른 답안: 투자→프로세스안내→본인인증방식
 - 오류 답안 : 대출 → 서류확인
 - ➔ 올바른 Intention과의 거리가 매우 멀
- Ensemble을 “모르는 문제” 를 판별하는데 이용
 - 예) 1번 머신: 투자 → 프로세스 안내
 - 2번 머신: 대출 → 서류 확인
 - 3번 머신: “계산 중”
 - ➔ “모르는 문제” 로 결론

풀고자 하는 문제: Attention Model

- Question Answering:

- 철수는 사과나 배를 먹었다.
철수는 사과를 먹지 않았다.
철수가 먹은 것은 무엇인가?
- 철수는 사과와 배를 먹었다.
철수는 감을 먹었을까?
- 제 계좌에 3천원이 있고,
와이프 계좌엔 5천원이 있는데
이걸 모두 제 계좌로 옮기고 싶어요.
그럼 제 계좌엔 얼마가 있게 되는것인가요?

Q: how old was monica lewinsky during the affair ?

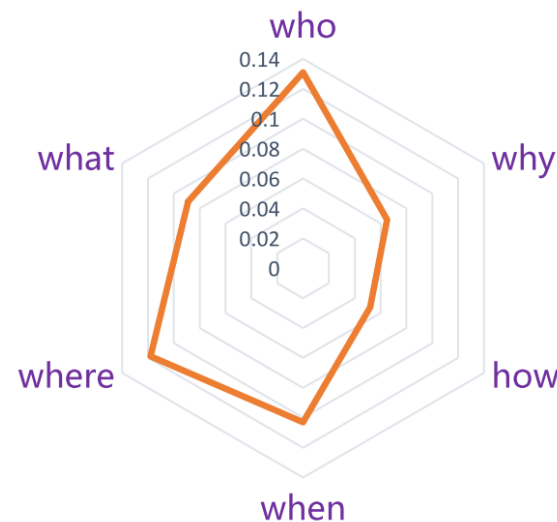
OARNN:

Monica Samille Lewinsky (born July 23 , 1973) is an American woman with whom United States President Bill Clinton admitted to having had an `` improper relationship '' while she worked at the White House in 1995 and 1996 .

IARNN-CONTEXT:

Monica Samille Lewinsky (born July 23 , 1973) is an American woman with whom United States President Bill Clinton admitted to having had an `` improper relationship '' while she worked at the White House in 1995 and 1996 .

B Wang et al, 2016. 8. 7



풀고자 하는 문제: 대화의 단절

- 대화의 단절
new line이 끝이 아니다
- Q: 제가 필요한 것은 전화번호
- A: 010-8282-8282 입니다.
- Q: 보다 팩스번호예요