

Revenue growth divisions.

TYU division

FRT division

Projected sales of main products in 2013

Natural Language Generation on Team Competition

MLP LAB 임경태

	TYU division			FRT division		
GHT	254	550	254	274	154	415
RDW	650	320	754	273	825	154
TRG	241	450	144	364	954	174
RTG	254	650	874	657	125	274
WEF	784	145	124	752	741	759
UDT	459	704	654	341	741	540



Contents

01 자연어 생성 task 소개

02 경진대회 소개

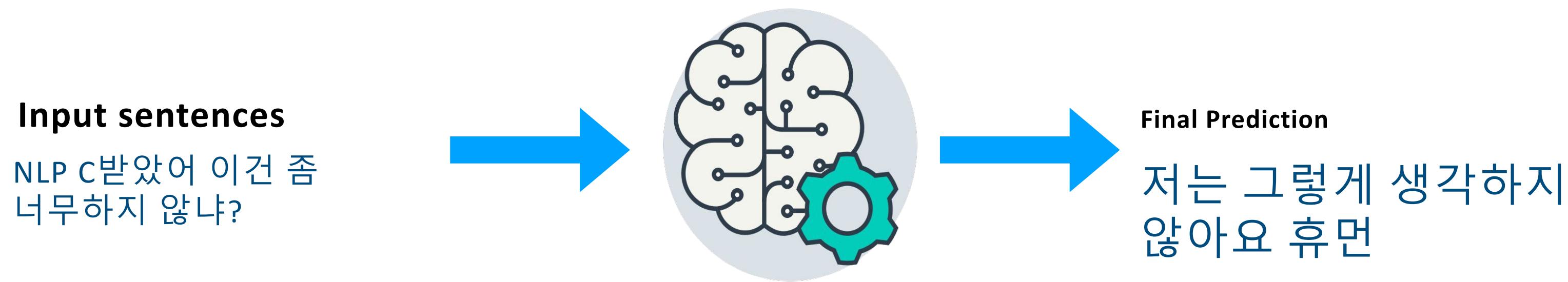
03 힌트



01. NLG (Chatbot)

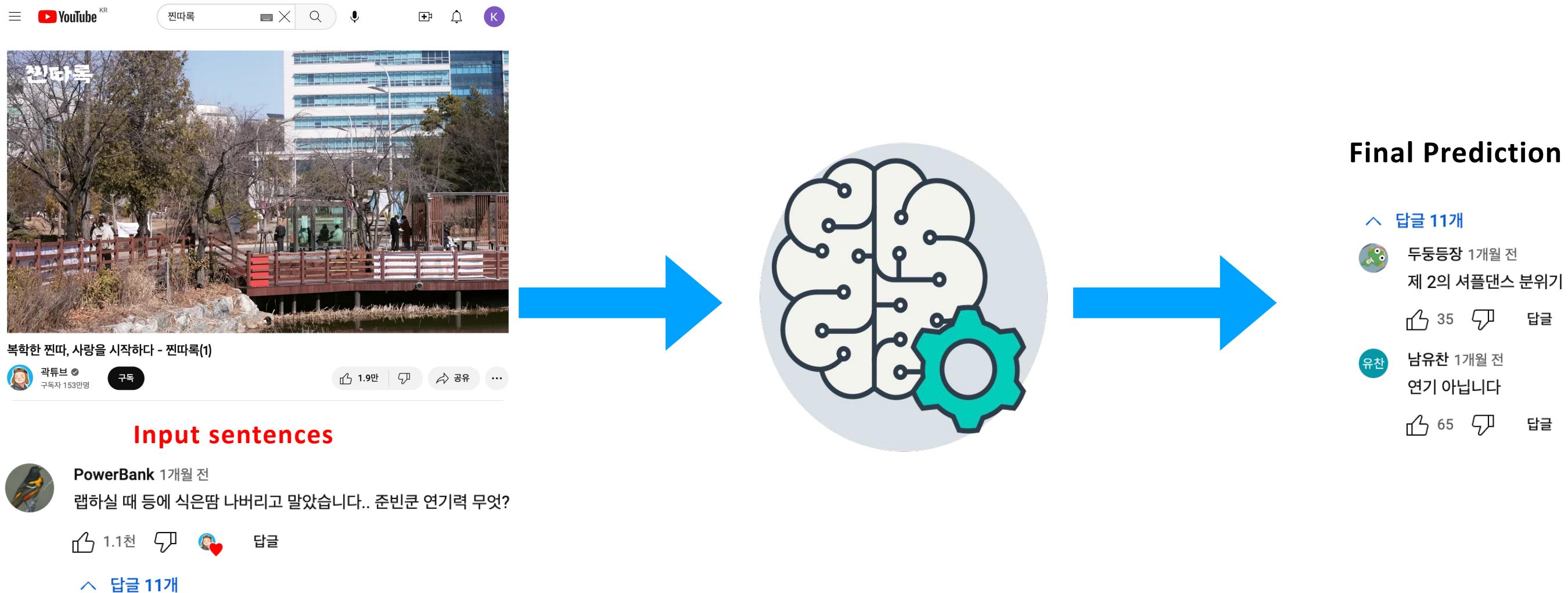
Natural language Generation (NLG)

- 자연어를 생성하는 Task를 의미하며 대표적으로 Chatbot이 있음



Natural language Generation (NLG) 예제!

- 자연어를 생성하는 Task를 의미하며 대표적으로 Chatbot이 있음
- 주어진 sequence to sequence를 이용해 유튜브 댓글 bot을 만들자!



Natural language Generation (NLG) 예제!

- 자연어를 생성하는 Task를 의미하며 대표적으로 Chatbot이 있음
- 주어진 sequence to sequence를 이용해 유튜브 댓글 chatbot을 만들자!

The image consists of three screenshots from the Google Cloud API dashboard:

- Screenshot 1: API 및 서비스**
Shows the main API dashboard with the "YouTube Data API v3" selected under "API 선택".
 - 1 사용자 인증 정보 유형**: A step-by-step guide for OAuth consent screen configuration.
 - 어떤 API를 사용 중이신가요?**: A list of available APIs categorized by service provider.
 - 액세스할 데이터는 무엇인가요?**: A section about scopes and permissions.
 - 선택한 기간에는 데이터가 없습니다.**: A note indicating no traffic data is present.
- Screenshot 2: 사용자 인증 정보**
A continuation of the OAuth configuration process, showing the "YouTube Data API v3" service details.
- Screenshot 3: API/서비스 세부정보**
Detailed information for the "YouTube Data API v3" service, including its provider (Google), service name (youtube.googleapis.com), and usage status.

Natural language Generation (NLG) 예제!

- 자연어를 생성하는 Task를 의미하며 대표적으로 Chatbot이 있음
- 주어진 sequence to sequence를 이용해 유튜브 댓글 chatbot을 만들자!



The screenshot shows a browser window with a YouTube video playing. The video title is "인공지능? 머신러닝? 딥러닝? 그 차이". Below the video, the URL is "youtube.com/watch?v=EMMkLEXN09I". To the right of the browser, there is a Jupyter Notebook cell containing Python code. The code uses the Google API client library to interact with the YouTube API. It imports necessary modules, defines a developer key and video ID, creates a comment thread list, and builds an API object to execute the comment threads list request.

```
[ ] from googleapiclient.discovery import build

import warnings # 경고창 무시
warnings.filterwarnings('ignore')

[ ] my_key = "AIzaSyC1I1yYG7Vx2_bMdJVe-1S7JFAJCeJ4aWU"
video_id = "EMMkLEXN09I"
comments = list()
api_obj = build('youtube', 'v3', developerKey=my_key)
response = api_obj.commentThreads().list(part='snippet,replies', videoId=video_id, maxResults=100).execute()
```



The screenshot shows a Jupyter Notebook cell displaying a DataFrame named 'df' containing YouTube comments. The columns are labeled 0, 1, 2, and 3. The data consists of five rows of comments, each with a timestamp and a user ID.

	0	1	2	3
0	연기쩐다ㅋ ALALALCNCN KOKIO 2023-05-21T16:10:47Z 0			
1	이 영상도 빠니보틀이 연출하고 대본 썼을까 그럼 기대되는 구나 상쾌한발사 2023-05-21T02:22:19Z 0			
2	향마력 딸린다ㅋㅋ 손수익 2023-05-20T14:14:14Z 0			
3	보는데 힘겨웠다. 대단하다.핥 리블링걸 2023-05-19T12:09:53Z 0			
4	빠니보틀은 계속 연출해서 작품 만들어주세요 너무 좋음 Hyeon 2023-05-19T06:33:59Z 0			

02. 천하제일 자연어 생성대회

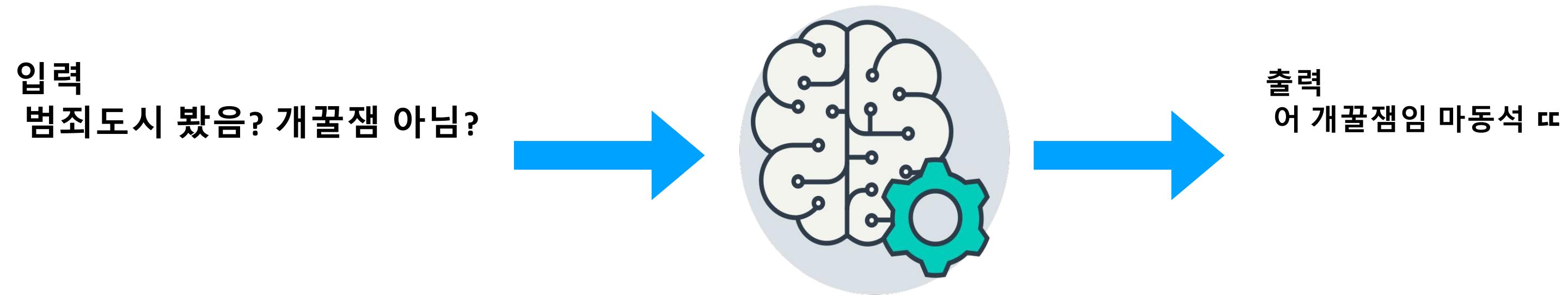
대회방법

- 3인 1조로 여가생활 관련 대화 생성모델 대회 참전
 - 목표: 지금까지 배운 내용을 토대로 2시간 동안 seq2seq 챗봇 모델을 강화하라!
 - 상품: 1등은 대학원 입학 우선권 dream(노예행 티켓 개꿀) 은 장난이고 고가의 미끼 제품들 준비



천하제일 여가생활 챗봇 구축

- 구현 예제

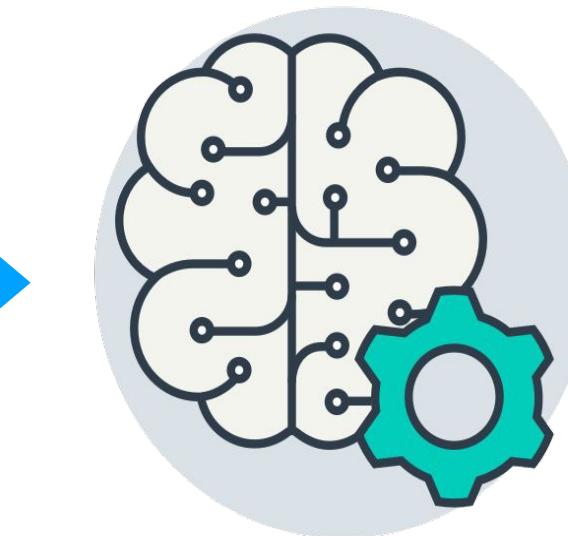


천하제일 여가생활 챗봇 구축

- 팀구성 예제

전처리

- 대화형 데이터 수집
- 대화형 데이터 전처리
- 대화형 데이터 가공



모델러

- 모델이 일단 돌아가게 만들어야 함
- 모델의 학습을 빠르게
- 모델의 성능을 좋게

후처리

- 모델의 출력물을 가공
- 실제 대화하는 느낌이 나도록 룰 기반 변경 혹은 예측 시 특정 단어 선택 알고리즘 구현
- generator() 함수 변경
- Github의 코드를 관리 및 운영

진행방식 (팀구성)

- 팀구성
 - 1) 모델러: 빠른 모델 (병렬처리..), 효율적인 모델, 좋은 성능을 보이는 모델 구현
 - 2) 전처리: 외부 학습데이터 수집 및 가공
 - 3) 후처리: 생성된 문장에서 규칙을 찾아 최대한 문맥에 맞도록 수정 및 github관리

진행방식 (평가)

- SSA(Sensible and Specificity Average) 사람이 직접 보고 평가
 - SSA는 사람이 직접 자율 발화 모델과 대화를 하며 점수를 평가하는 것인데, Sensibleness와 Specificity의 두 가지 항목 외엔 점수 측정. Sensibleness은 사람의 말에 대한 봇의 답변이 합리적인지를, Specificity는 해당 답변이 단답형이 아닌 구체적인 답변인지를 나타냄.

학습데이터 참조

<https://www.aihub.or.kr/aihubdata/data/view.do?currMenu=115&topMenu=100&aihubDataSe=realm&dataSetSn=114>

<https://www.aihub.or.kr/aihubdata/data/view.do?currMenu=115&topMenu=100&aihubDataSe=realm&dataSetSn=543>

감사합니다.