



KLUE 한줄 요약 프로젝트 오리엔테이션

13기 금지현 13기 기다연 13기 박주영
12기 이가영 13기 홍석민



오티 진행 순서



01

프로젝트 개요 (지현)

- 주제
- 요약 태스크 설명
- 프로젝트 일정 설명

02

스터디 개요(주영)

- 논문별 주제 설명
- 논문 스터디 진행 방식
- 논문 스터디 일정
- 스터디 블로그 정리 설명

03

영상 발표 (지현, 주영)

- Machine Translation Survey (vol1) : Background (주영)
- Transformer to T5 (지현)

04

기타

스터디 발표자 선정 /
회의 일정 정하기 /
주차별 회의 서기 선정

05

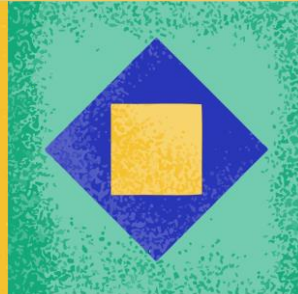
친목

팀명 만들기^^;
말놀이..?

02.

Study introduction

9월 16일부터 11월 25일까지 총 11주에 걸쳐
9개의 논문 스터디 진행



PAPER TABLE



	MODEL	Multi-document	metric
Title 1	Attention Is All You Need	Bottom-Up Abstractive Summarization	ROUGE: A Package for Automatic Evaluation of Summaries
Title 2	Fine-tune BERT for Extractive Summarization	TextRank: Bringing order into texts	BERTScore: Evaluating Text Generation with BERT
title 3	BertSum: Text Summarization with Pretrained Encoders	Few-Shot Learning for Opinion Summarization	
Title 4	BART: Denoising Sequence-to-Sequence Pre-training for Natural Language Generation, Translation, and Comprehension		

MODEL



Attention is all you need

RNN, LSTM 등에서 나타나는 순차적인 계산의 제약(입력과 출력 사이의 의존성)을 해결하기 위해 attention 메커니즘에만 기초한 **Transformer**를 제안



BertSum: Text Summarization with Pretrained Encoders

여러 sentence를 하나의 인풋으로 넣어주기 위해 제안된 BERTSUM과 그 외 layer를 변형 시켜 생성된 ext, abs 방식의 모델을 제안



Fine-tune BERT for Extractive Summarization

지문에 있는 단어와 문장을 가져와 그대로 요약하는 방식인 Extractive Summary를 BERT를 활용해 테스트



BART: Denoising Sequence-to-Sequence Pre-training for Natural Language Generation, Translation, and Comprehension

손상된 text를 복구하도록 사전 학습되어 noise로부터 자유로운 BART 소개 및 여러 noising 기법 평가

Multi-document



Bottom-Up Abstractive Summarization

Abstractive summarization의
한계점인 content selection을 다루는
간단한 기술을 제안하고 텍스트 요약에
있어 bottom-up attention 단계에 적용



TextRank: Bringing order into texts

각 정점의 정보만을 고려하지 않고 전체
그래프의 글로벌 정보를 재귀적으로
계산하여 정점의 중요도를 결정하는
방법인 graph-based ranking
알고리즘을 소개



Few-Shot Learning for Opinion Summarization

의견 요약과 같은 방식은 매우 중요하나
supervised 데이터가 부족하다는
한계가 있었는데 적은 수의 요약만으로도
여러 속성의 텍스트 요약이 가능하다는
봇스트랩 생성에 충분함을 제안

metric



Z

ROUGE: A Package for Automatic
Evaluation of Summaries

텍스트 자동 요약, 기계 번역 등 자연어
생성 모델의 성능을 평가하기 위한
지표이며, 모델이 생성한 요약본 혹은
번역본을 사람이 미리 만들어 놓은
참조본과 대조해 성능 점수를 계산



N

BERTScore: Evaluating Text
Generation with BERT

참조 문장과 생성 문장을 각각 BERT에
태워 문맥 벡터(Contextual
embedding)를 구한 후 코사인 유사도를
이용



TIMELINE

09.23

- BertSum: Text Summarization with Pretrained Encoders
- BART: Denoising Sequence-to-Sequence Pre-training for Natural Language Generation, Translation, and Comprehension

11.11

- TextRank: Bringing order into texts
- Few-Shot Learning for Opinion Summarization

09.16

- Attention is All You Need
- Fine-tune BERT for Extractive Summarization

09.30

- Bottom-Up Abstractive Summarization

11.25

- ROUGE: A Package for Automatic Evaluation of Summaries
- BERTScore: Evaluating Text Generation with BERT

How?



presenter

논문 당 발표자 2인을 선정해
해당 발표자들은 티스토리
블로그에 발표 내용을 정리해
업로드



reader

나머지 인원들은 각자
자율적으로 공부

