



국민청원데이터터

확률 계산 AI

지윤혁 | 강현빈
김민주 | 양서정

INDEX

| BLUEPRINT |

주제 및 기획 방향

| AI |

데이터 처리 및 주요 모델 구현

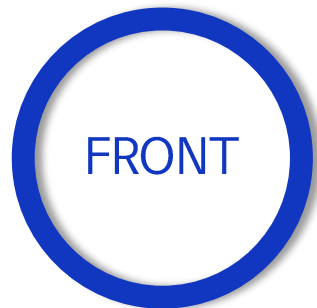
| TECH |

프론트, 백

| SIMULATE |

구현과 시현

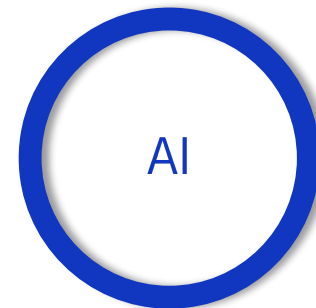
TEAM 2



HTML	김민주 양서정
CSS	김민주 양서정
JS	강현빈



DB 연동	김민주 양서정 강현빈
FRONT 연동	김민주 양서정



논문정리	지윤혁
크롤링	지윤혁
모델링	지윤혁
AI 연동	지윤혁 강현빈

REALIZED

	0506	0507	0508	0509	0510	0511	0512
AI	Transformer, BERT 리뷰	GPT, ELMO 리뷰	국민동의청원 사이트 크롤링	파파고 영어번역 크롤링	LSTM, Transfomer 모델 학습	GPT, BERT 모델 학습	Project발표
Front-end		FRONT 1차	FRONT 재시작 홈, Input	Log In, Result	Result, Board	Board	
Back-end			DB 연동	DB, AI 연동	AI 연동		

PLANNED

	0506	0507	0508	0509	0510	0511	0512
AI	최신 논문 수집	논문 리뷰	Model 학습 및 구현				Project발표
Front-end				Web Front-end 구현			
Back-end					Web Back-end 구현		

BLUEPRINT



INTRO



<https://news.kbs.co.kr/news/view.do?ncd=4105372>

GOAL

첫째, 무의미한 글 가십화 또는 남발 방지
둘째, 국민 청원 사이트 유저들에게 채택율 예측치 제공



국민과의 소통, 직접 민주주의 실현이라는 본질적 목표에 초점을 맞출 수 있게 지원

PLANNED

개인이 생성한 텍스트 기반 우울 여부 판독

REALIZED

국민 청원 데이터 기반 청원 채택율 예측

사용하고자 했던 자연어 분류 모델 학습 및 활용

구상한 화면 설계 완료

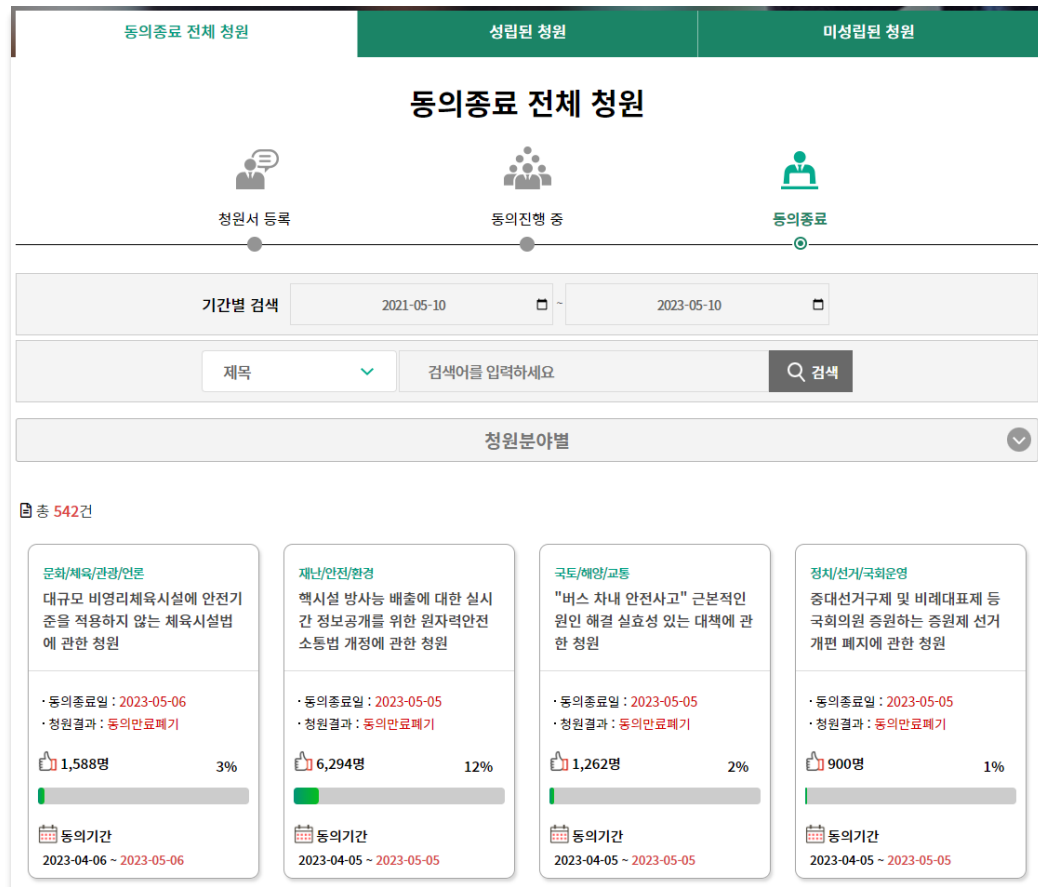
동적인 화면 구상

정적인 화면으로 대체

AI



DATA | DATA Processing | MODEL | RESULT



papago 웹사이트 번역 문서 번역 GYM 사전

한국어 감지

영어

A petition for legal regulations to prevent indiscriminate use and abuse of AI image generators

The development of AI image generators has great potential to bring new innovations to the content industry. However, the risks and negative consequences cannot be overlooked. Without regulations and ethical guidelines, AI image generators can harm individuals and society as a whole. AI image generators can easily be used to spread false news, information, and defame real people. This not only causes confusion throughout society, but also violates individual rights. In addition, the production and commercial use of images that infringe on copyright or portrait rights can cause financial damage to the original author. Some might argue that regulation hinders development, but it is not. Regulations and ethical guidelines promote a win-win relationship between AI image generators and creators, enabling the content industry to evolve in a responsible and respectful way for everyone involved. In addition, ethical AI development can increase the reliability of AI technology and escape controversy over data unauthorized learning such as copyright and portrait rights. Therefore, it is important to create an environment in which creators can actively accept AI as a tool and a social perception that everyone can sympathize with. We saw the negative impact of AI image generators on society and the content industry. AI image generator regulatory petitions are necessary measures to promote ethical AI and protect creators' intellectual property rights. It's not about stifling innovation, it's about developing AI in a responsible and ethical way that benefits everyone. The conclusion cannot ignore the potential damage caused by unregulated AI image generators. We need to take steps to protect individual rights and promote the development of ethical AI that benefits everyone. Please participate in the AI generator image regulatory petition support for the future of the content industry.

962 / 5000

번역하기

자동한성

어 퍼티션 포어 리걸 레귤레이션즈 투 프리렌트 인디스크리미넨트 유스 언드 어부스 어브 에이아이 이미지 제너레이터즈 더 더덜립먼트 어브 에이아이 이미지 제너레이터즈 헤즈 그레이트 퍼텐셜 투 브링 뉴 이노베이션스 투 더 콘텐츠 인더스트리 ...더보기

번역 수정 | 번역 평가

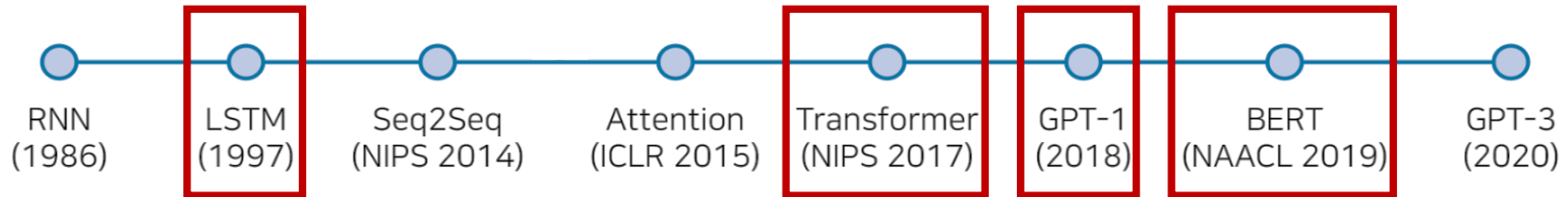
한국어 처리

- ▷ 영문, 숫자, 한글을 제외한 특수문자 제거
- ▷ 형태소 분석
- ▷ 불용어 처리
- ▷ 각각의 어절을 분리
- ▷ 표준화
- ▷ 정규화
- ▷ 토큰나이징
- ▷ 패딩

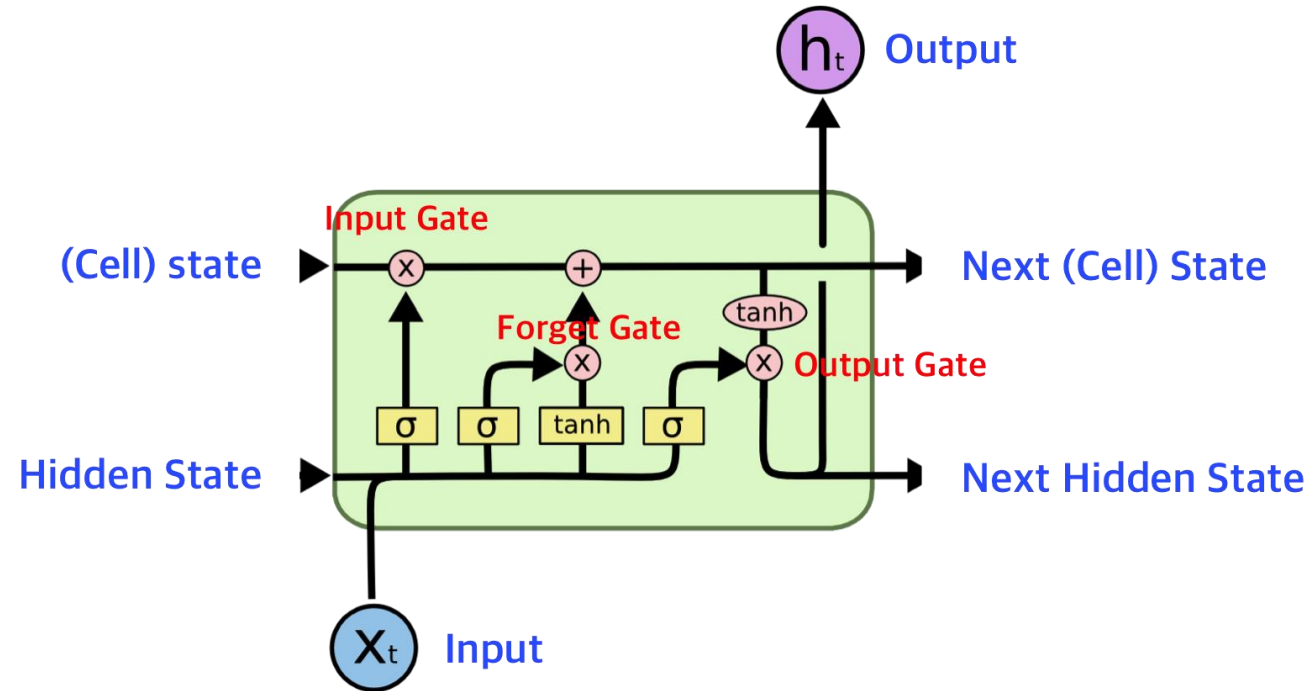
영어 처리

- ▷ 소문자로 통일
- ▷ 특수문자 및 숫자 제거
- ▷ 불용어 처리
- ▷ 토큰나이징
- ▷ 패딩

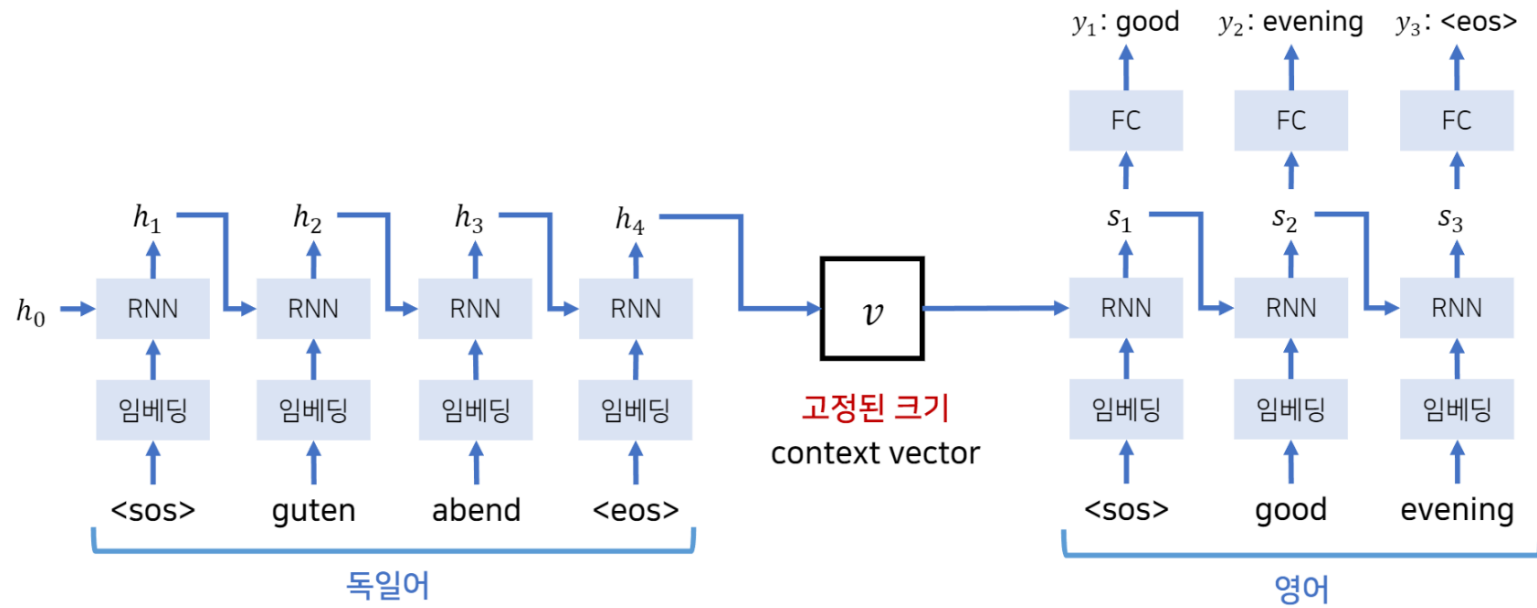
Machine Translation



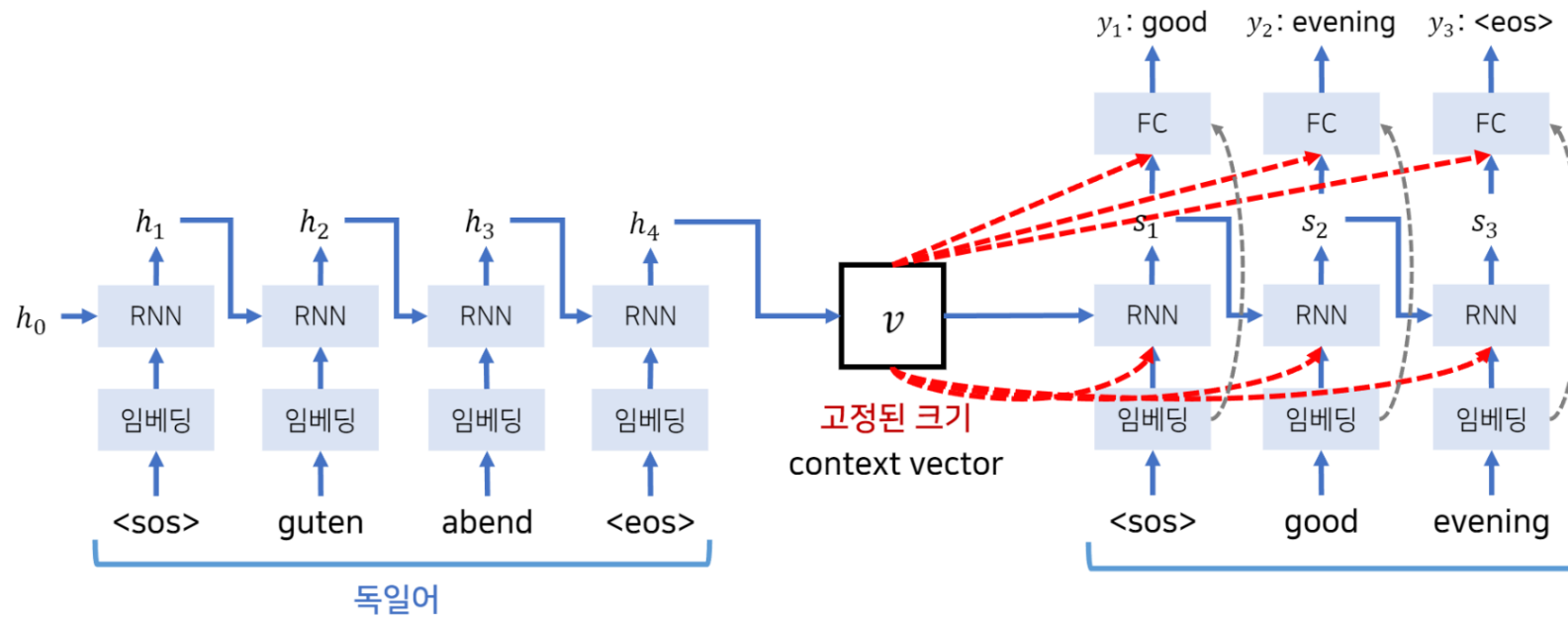
LSTM : Long Short Term Memory



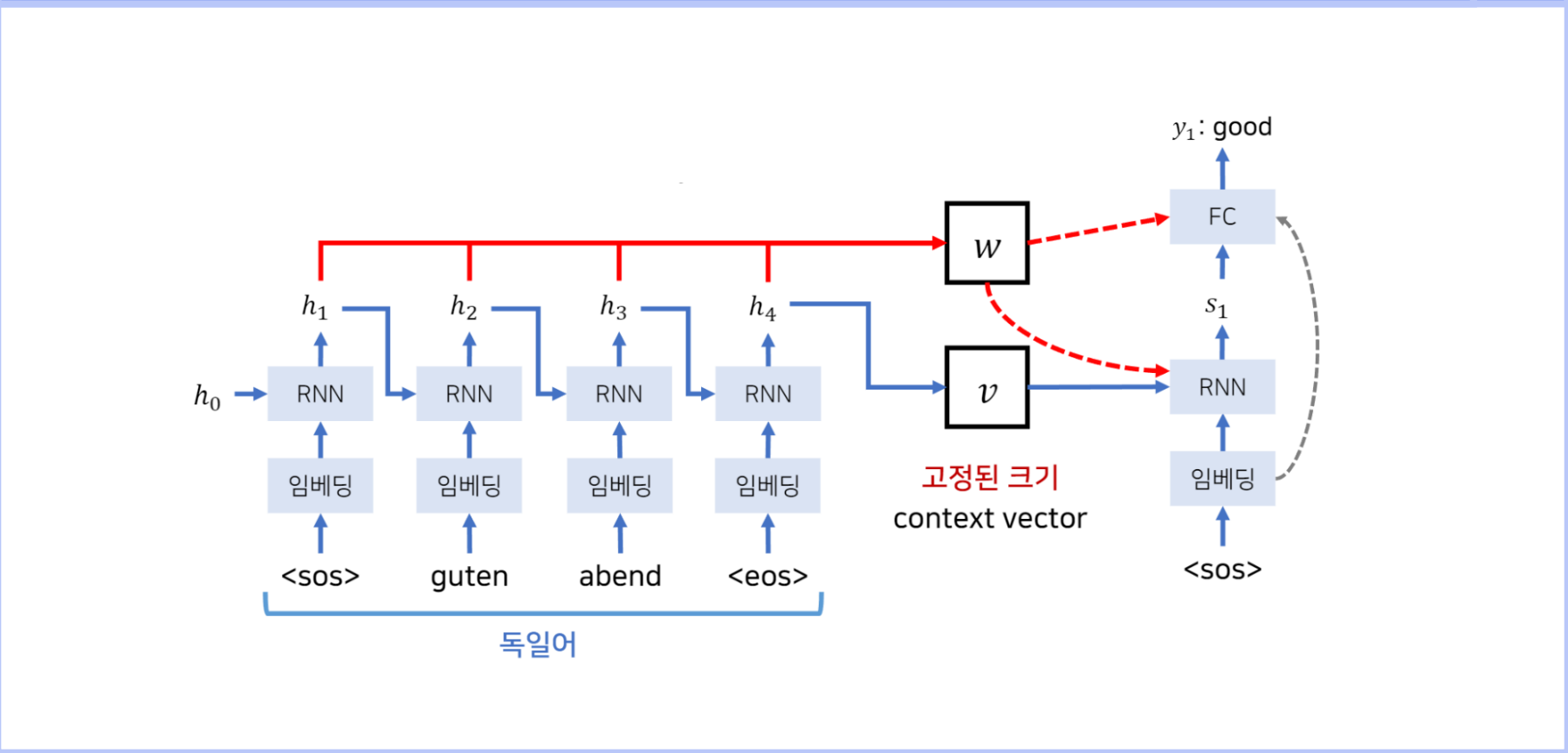
Seq2Seq : Sequence to Sequence Learning with Neural Networks | 2014



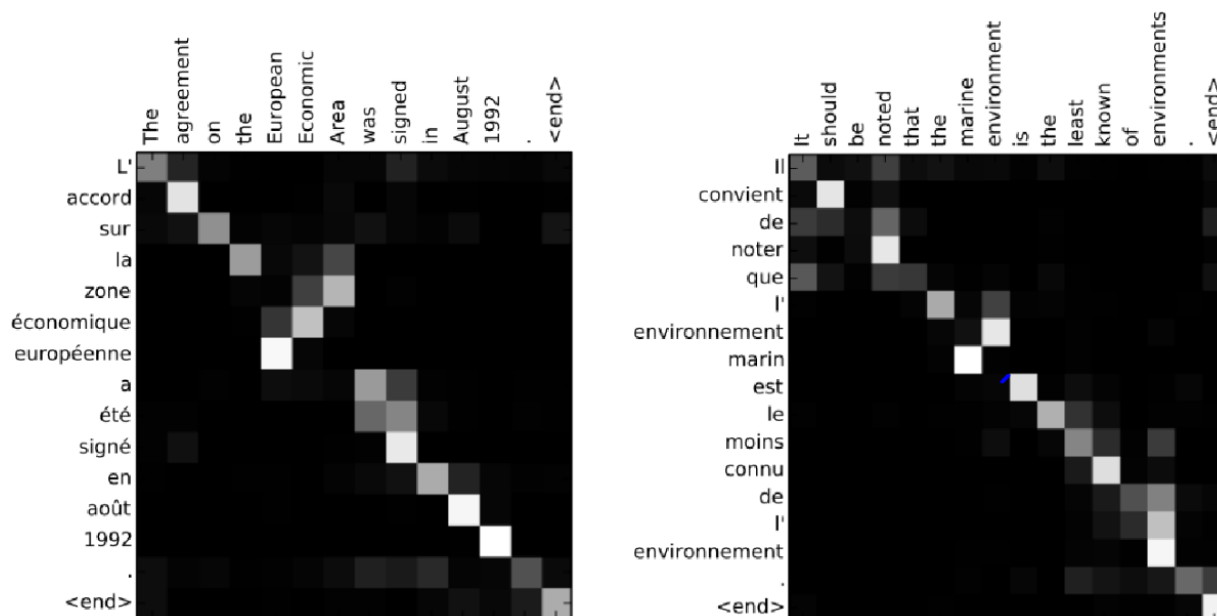
Seq2Seq : Sequence to Sequence Learning with Neural Networks | 2014



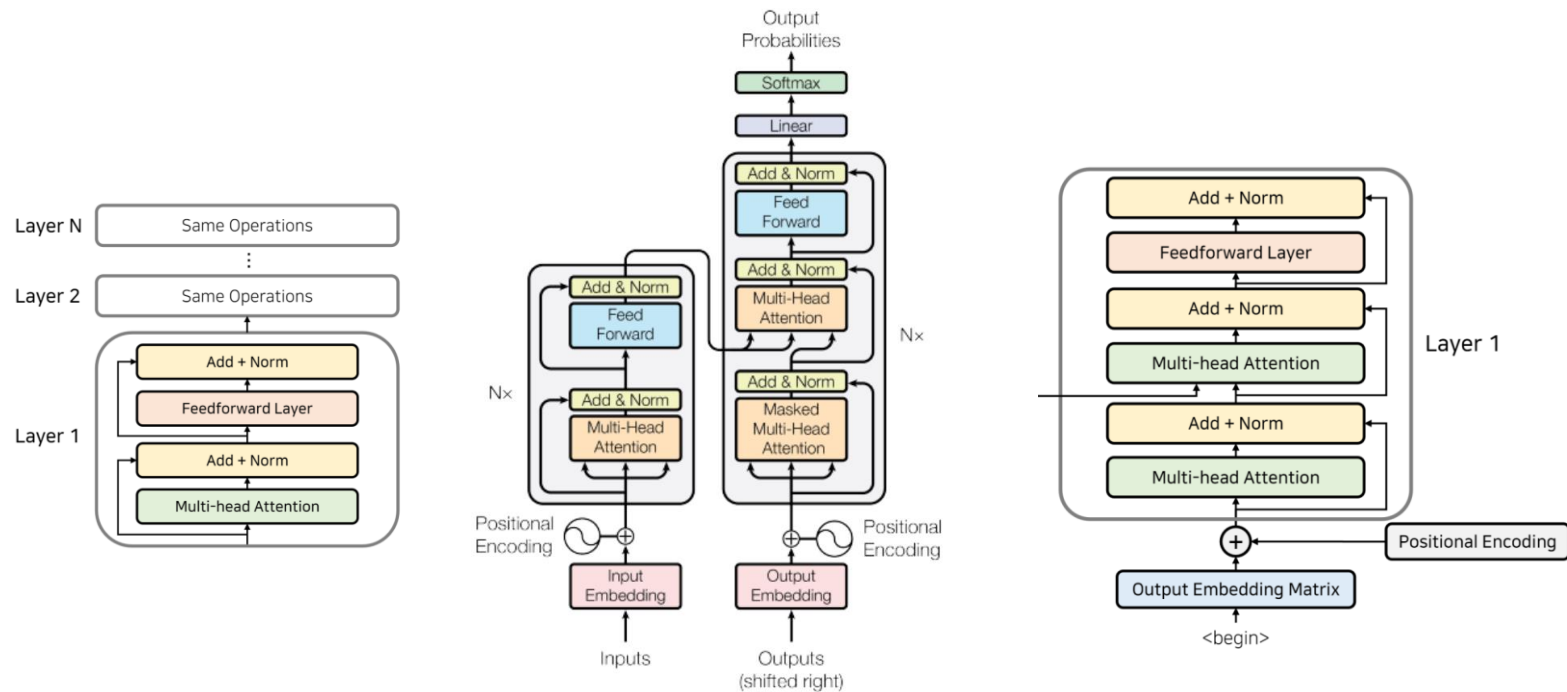
Seq2Seq With Attention : Sequence to Sequence Learning with Neural Networks | 2014



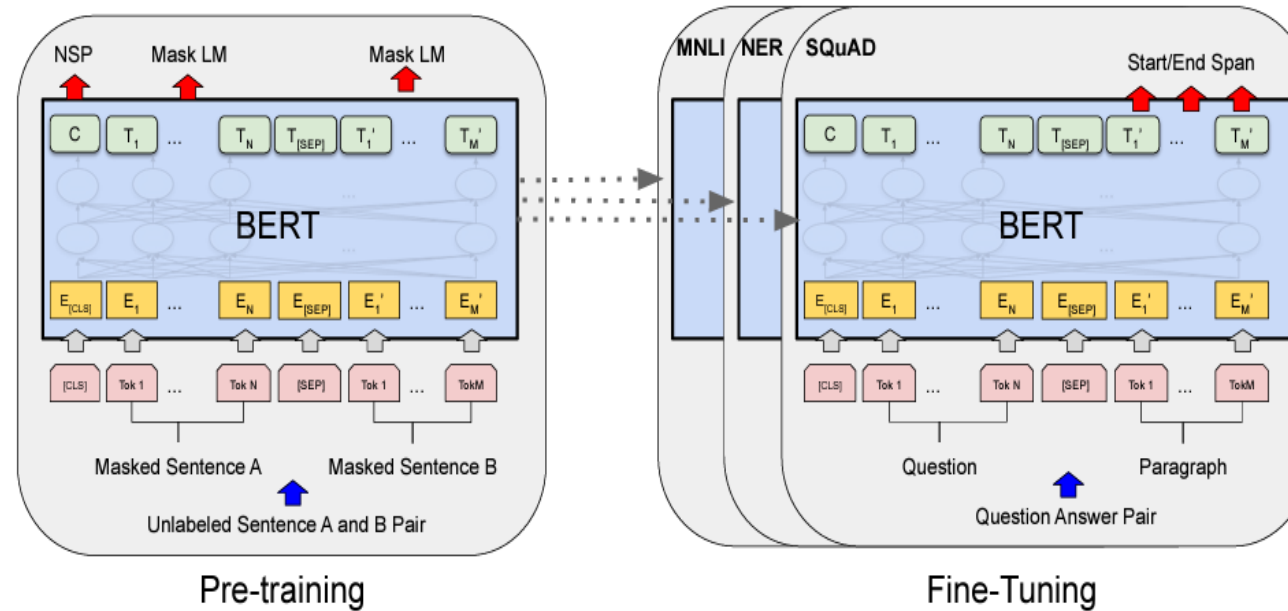
Attention Map



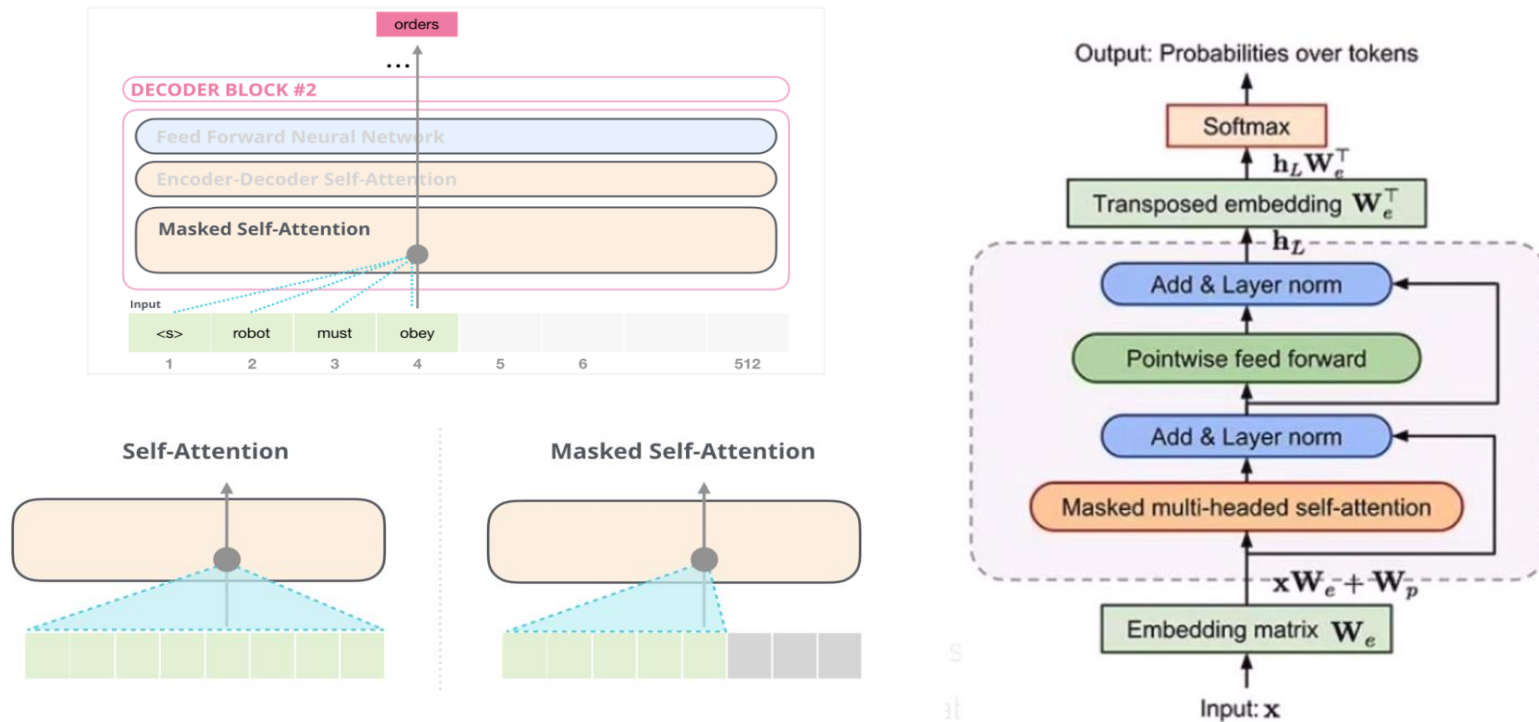
Transformer : Attention Is All You Need | 2017 |



BERT : Pre-training of Deep Bidirectional Transformers for Language Understanding | 2018 |

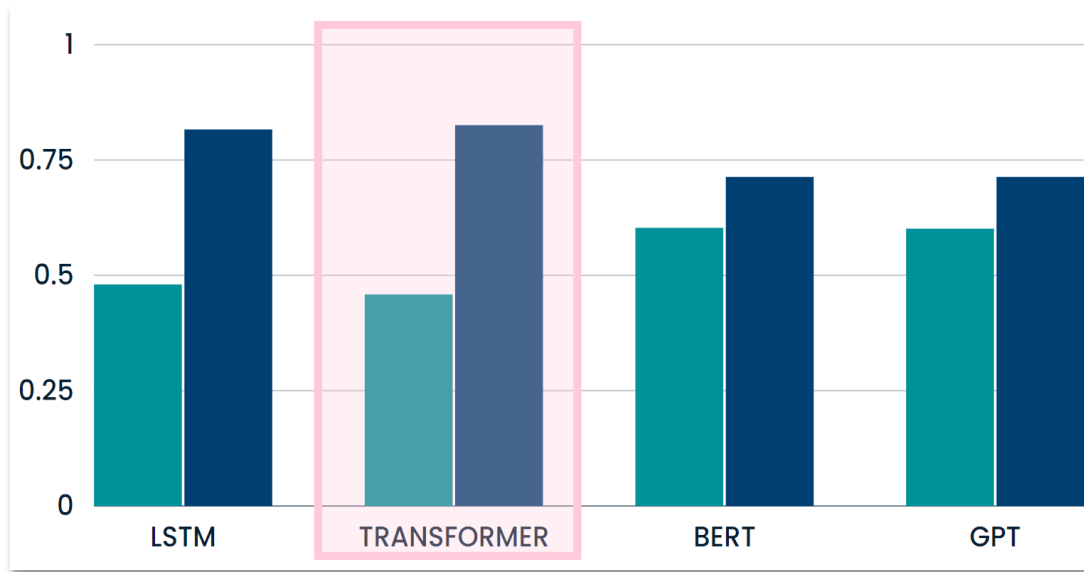


GPT : Improving Language Understanding by Generative Pre-Training | 2018

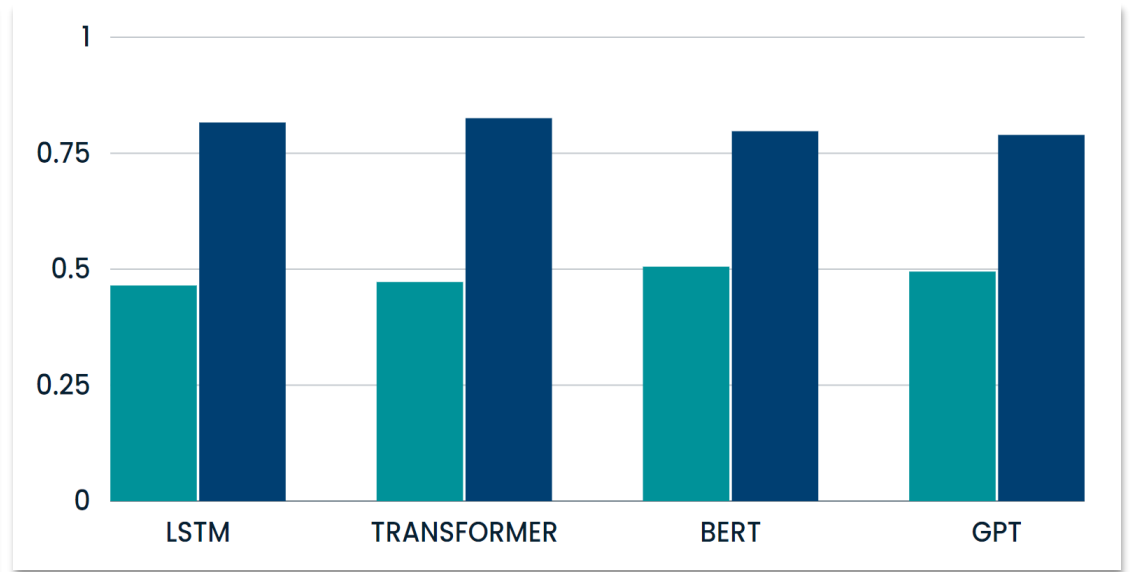


RESULTS

KOREAN

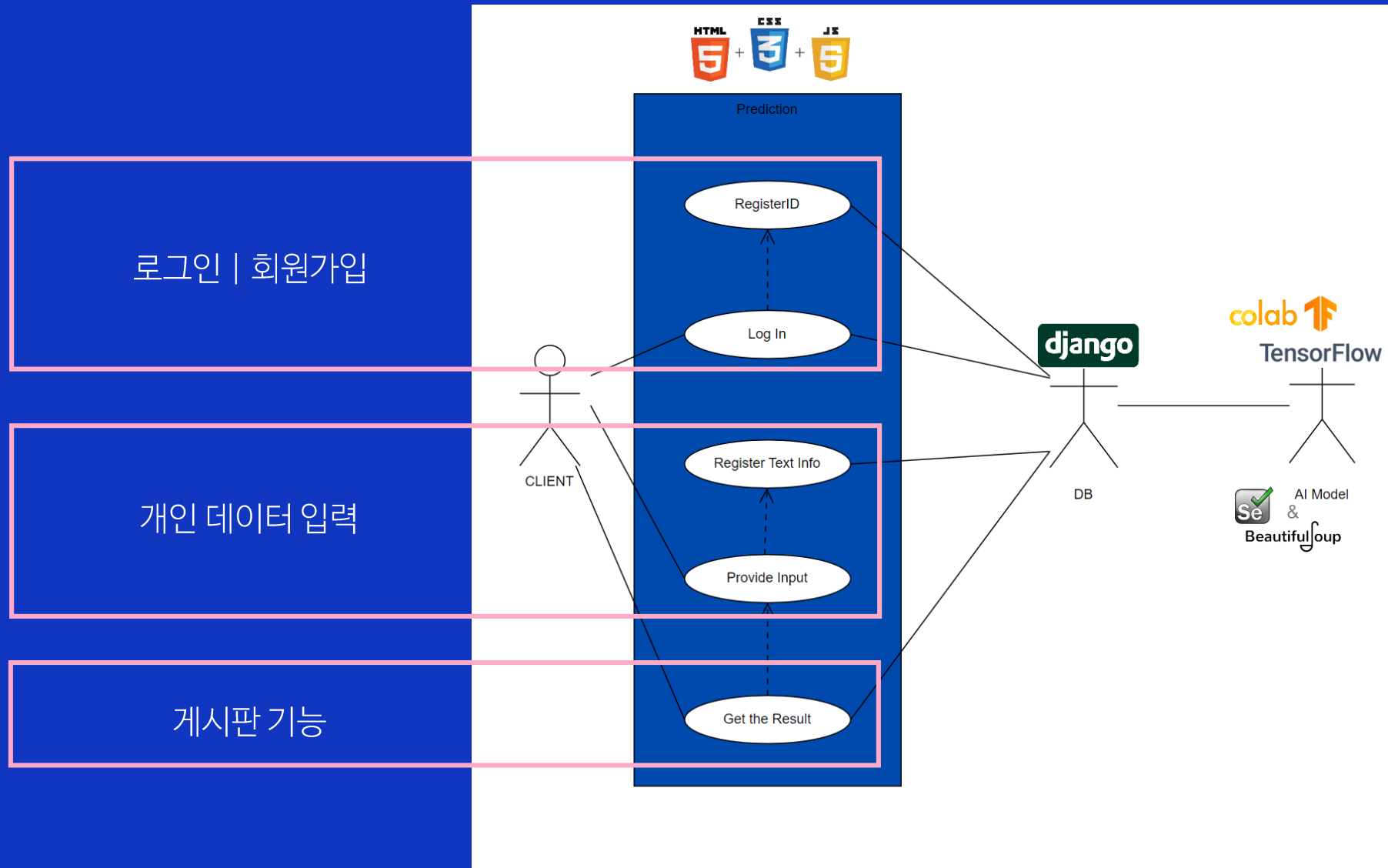


ENGLISH



TECH





SIMULATE



REFLECTION



COMMENTARY

| 지윤혁 |

이전까지 진행했었던 프로젝트와는 다르게 시간적 여유가 없었기 때문에 어떠한 데이터로 어떠한 모델을 선택해서 AI 프로젝트를 진행할 지 걱정이 많았습니다. 여러 고민 끝에 프로젝트의 목표를 기존 수업시간에 배웠던 부분에서 확장하여 여러 논문을 참고하고 실제로 데이터에 적용해보고, 성능을 비교해볼 수 있는 방향으로 진행하고자 하는 의견을 팀원들이 지지해줘서 프로젝트를 잘 마무리할 수 있었던 것 같습니다. 연휴 포함 대략 2주 가량 진행된 프로젝트에 밤낮으로 열심히 고생해주고, 저의 의견을 지지해주셨던 모든 팀원들께 감사의 말씀을 전하고 싶습니다. 정말 고생 많으셨습니다.

| 강현빈 |

웹 경험자가 1명도 없던 조라, 처음 계획했던 것에서 적지 않은 부분이 바뀌어 아쉬웠습니다. 백엔드 쪽을 주로 맡았는데, 수업 때 배운 MTV패턴을 기억하며 흐름을 따라가고자 했습니다. 하지만 회원가입 및 로그인 기능을 구현하는 과정에서 DB와 연동할 때, 가장 많은 어려움을 마주해서 힘들었던 것 같습니다. 하지만 에러를 마주할 때마다 원인이 무엇이고 해결방법은 무엇인지 기록해 갔고, 뒤로 갈수록 디버깅에 걸린 시간이 단축되는 것에 서로 나름의 성취감을 느끼기도 했었습니다. 크게 깨달은 것은 백엔드 담당이 탄탄해야 전체적인 진행이 원활하겠구나였는데 제 실력이 부족해 미안했던 마음도 많이 컸습니다. 계속 공부해야할 것 같아요. 다들 너무 고생한 2주였습니다..!

| 양서정 |

Web 경험자가 단 한 명도 없는 조에서 고군분투 했던 2주였습니다. 극히 단순한 메커니즘을 가진 웹 사이트지만, 프론트 부터 백까지 맨 땅에 헤딩하듯 구현해가는 과정에서 많은 교훈을 얻었습니다. 동적이고 사용자 친화적인 프런트를 구현해보고 싶었고, 더 복잡한 메커니즘을 가진 서비스의 백도 경험해보고 싶습니다. 조금 더 거시적인 시각에서는 유스 케이스를 촘촘하게, 그리고 Views.py를 통해 조금 더 구체화 한 후에 프론트 설계를 진행하는 것이 효율적일 수 있겠다는 생각이 들었습니다. 물론 초보 개발자의 감상입니다.

| 김민주 |

Web 경험자가 없어서 처음에는 막막했지만, 오히려 그 덕분에 프론트부터 백까지 모두 경험해볼 수 있는 의미있는 시간이었습니다. 프로젝트를 진행하면서 각자의 역할을 잘 수행하는 것도 중요하지만, 각 역할 사이의 소통도 굉장히 중요하다는 것도 다시 느낄 수 있었습니다. 프로젝트 기간이 짧아서 더 많은 기능을 구현해보지 못한 점이 아쉬웠습니다. 하지만 이번 경험을 바탕으로 이후에 웹프젝트가 진행되면 더 수월하게 진행해서 팀에 더 도움이 되는 사람이 될 수 있을 것 같습니다. 고생 많으셨습니다!

APPENDIX



| Transformer 논문 정리 | <https://burly-freckle-bbd.notion.site/Transformer-20b0dc3c182d4c22a4717d75393362f5>

| BERT 논문 정리 | <https://burly-freckle-bbd.notion.site/BERT-a70dd3d299c04525af811afa887d5084>

| GPT 논문 정리 | <https://burly-freckle-bbd.notion.site/GPT-1b0fbb2f257c4ffa9be289c9b9e8895f>

| ELMO 논문 정리 | <https://burly-freckle-bbd.notion.site/ELMo-4897328e1b5143d0a9c269f88751024c>

APPENDIX

KOREAN MODEL

| **LSTM** | <https://colab.research.google.com/drive/1NwJjMSeyvNBZqkH42nJBjWqeeGUk0lD1?usp=sharing>

| **Transformer** | https://colab.research.google.com/drive/1KQlhSo_TM9XKRXgnbWMD101_5HYb6vKB?usp=sharing

| **BERT** | <https://colab.research.google.com/drive/1j5KtABC8bjtIJM9Wf1Tt9dBGODBgNHp?usp=sharing>

| **GPT** | <https://colab.research.google.com/drive/1evm04qfzZbPQL2oQzTTkB2ZfMDkz1U3m?usp=sharing>

APPENDIX

ENGLISH MODEL

| **LSTM** | <https://colab.research.google.com/drive/1Eoc9fY8Z27danufA4p2InsUOkvw4gpn8?usp=sharing>

| **Transformer** | <https://colab.research.google.com/drive/1z1nDOahG0fSLyST0kfyCdKstu71WwHkF?usp=sharing>

| **BERT** | <https://colab.research.google.com/drive/1R6O-u4t8aFndw7RGqAHJ7xTTg3G7ox8S?usp=sharing>

| **GPT** | https://colab.research.google.com/drive/1VZ9LGsXn_7YU1dpR8CUJYCOMiKEg1ak0?usp=sharing

THANK YOU

