

# 플리토 AI 번역 엔진

flitto.com과 앱에서 서비스 중인 엔진

언어쌍별 1천만쌍 이상의 데이터 학습

Transformer 모델 사용

9개 언어쌍 번역 가능

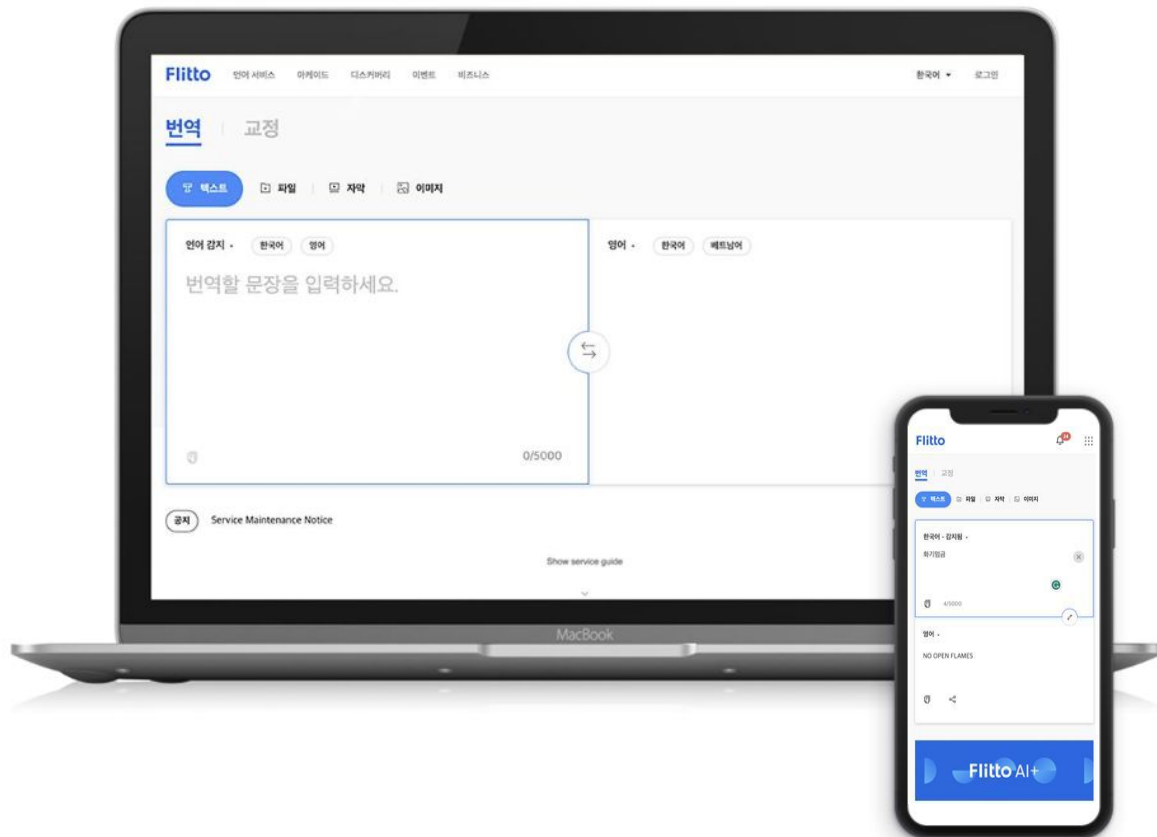
한국어 <-> 영어

한국어 <-> 일본어, 한국어 <-> 중국어

영어 <-> 일본어, 영어 <-> 중국어

영어 <-> 스페인어, 영어 <-> 독일어

영어 <-> 러시아어, 영어 <-> 프랑스어



## 플리토 AI 번역 엔진

### 관련 보도 자료

'사업모델 특례상장 1호' 플리토, 자체  
데이터 기술로 'AI 번역 엔진' 상용화

<https://www.etnews.com/20200601000300>

플리토 '기업 맞춤형 AI 번역' 출시

<https://www.venturesquare.net/824503>

# '사업모델 특례상장 1호' 플리토, 자체 데이터 기술로 'AI 번역 엔진' 상용화

발행일 : 2020-06-01 15:42 지면 : 2020-06-02  3면

173개국서 1000만명 이상 사용자 확보  
집단지성 플랫폼으로 방대한 DB 학습  
실시간으로 '산업 특화 번역' 차별화  
e커머스·여행·숙박 등 활용도 높아

## 플리토 AI 번역 엔진

### 프로젝트내 역할

데이터 수집, 전처리부터

모델 학습, API 개발,

On-premise 고객 지원까지

ML part 전과정 담당

## 사용 기술

데이터 수집: Python web crawling

전처리: SentencePiece, NER, tokenizer

모델 학습: PyTorch (fairseq)

API 개발: Flask / FastAPI

서버 배포: AWS EC2

# 플리토 AI 번역 엔진

## Performance

타사 번역 엔진(Papago, Google)과

대등한 번역 품질

사내 Linguist 대상 Blind test 결과

일부 언어 쌍에서 파파고/구글 번역보다

우수한 번역 품질을 보였음.

[원문]

Lime's withdrawal comes two years and seven months after the company first launched its operations in October 2019. Lime operated roughly 22,000 vehicles in South Korea at one point, according to the spokesperson. The company didn't say when it would return to South Korea, the only Asian country Lime operates in, but the spokesperson told TechCrunch Lime intends to continue working with the government and other operators to establish a regulatory framework that will allow for the long-term success of micromobility in the country.

[추천 번역 (사람이 한 번역)]

라임의 철수는 2019년 10월 회사가 처음 사업을 시작한 지 2년 7개월 만이다. 대변인에 따르면, 라임은 한국에서 한때 약 22,000대의 차량을 운영했다고 한다. 회사 측은 라임이 진출한 유일한 아시아 국가인 한국에 언제 돌아올지는 밝히지 않았으나, 대변인이 테크크런치에 언급하기로는 라임이 국내 마이크로모빌리티의 장기적인 성공을 위한 규제 프레임워크를 구축하기 위해 정부 및 다른 사업자들과 지속 협력해 나갈 생각이라고 밝혔다.

[플리토 번역기]

라임의 철수는 2019년 10월 회사가 처음 영업을 시작한 지 2년 7개월 만이다. 라임은 한때 한국에서 약 22,000대의 차량을 운영했다고 대변인은 전했다. 회사 측은 라임이 운영하는 유일한 아시아 국가인 한국에 언제 돌아올지는 밝히지 않았지만, 대변인은 테크크런치라임이 정부 및 다른 사업자들과 계속 협력해 국내 마이크로모빌리티가 장기적으로 성공할 수 있는 규제 틀을 구축할 생각이라고 밝혔다.

[타사 번역기]

라임의 철수는 2019년 10월 처음 사업을 시작한 지 2년 7개월 만이다. 대변인에 따르면 라임은 한 시점에 한국에서 약 22,000대의 차량을 운영했다. 회사는 Lime이 운영하는 유일한 아시아 국가인 한국에 언제 반환할지 밝히지 않았지만 대변인은 TechCrunch Lime이 정부 및 기타 사업자와 계속 협력하여 장기적으로 허용할 규제 프레임워크를 구축할 것이라고 말했습니다. 국가에서 마이크로 모빌리티의 기간 성공.

이 부분에서 두 번역기의 차이가 좀 더 확연하게 드러나는데요. 플리토 번역은 전체적으로 오역이나 누락이 없는 반면, 타사 번역에서는 몇 가지 단어가 해석에 혼선을 주는 듯합니다. 예를 들면, '한국에 언제 돌아올지는 밝히지 않았지만'의 문장에서 타사 번역은 '언제 반환할지 밝히지 않았지만'으로 번역되어 아리송한 의미가 돼 버립니다. 또, 'at one point'는 타사 번역의 '한 시점에'보다는 플리토가 번역한 '한때'가 훨씬 더 자연스럽습니다.

출처: [영어 번역기 뭘쓰세요? 성능 보고 고르세요!-국내 사용자 1위 번역기 vs. 플리토 번역 결과 전격 비교](#)

# 인공지능 번역 API

## B2B 고객 대상 API 솔루션 중

### 내부 서버간의 API 개발

### 부가 기능

한국어, 일본어로 번역시 높임말 지원

기업 고객별 용어집(Translation Memory)

Pivot Translation: 한국어 -> 영어 -> 일본어

## API 솔루션

플리토 번역 API는 플리토의 다양한 번역 서비스를 RESTful 형태의 API를 통해 빠르고 간편하게 이용하실 수 있는 서비스입니다.

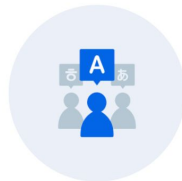
인공신경망 기반의 인공지능번역과 전세계 언어전문가들을 통해 실시간 번역되는 집단지성 번역을 통해 더욱 자연스러운 번역 결과를 얻을 수 있습니다.

## API 종류



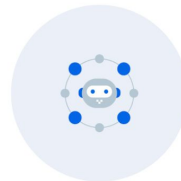
### 인공지능 번역 API

자체 개발 인공신경망 기술을 활용한 누적학습 2,000만 문장 이상의 고품질 인공지능 번역



### 집단지성 번역 API

전세계 300만명 이상의 언어전문가를 통한 보다 자연스러운 번역



### 기업 특화 번역 API

기업 전용 서버에 분야별 언어데이터 학습을 통해 만들어진 기업 환경에 특화된 인공지능번역

# 메신저 번역 챗봇

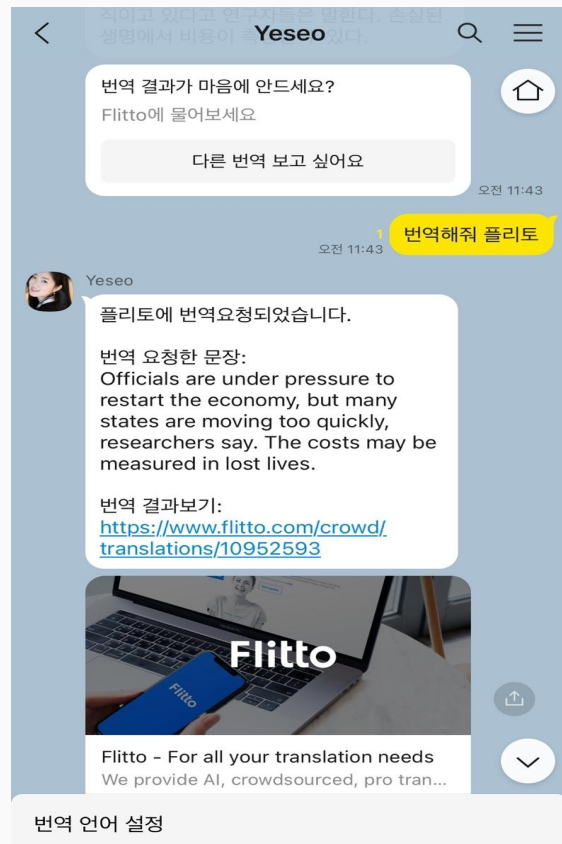
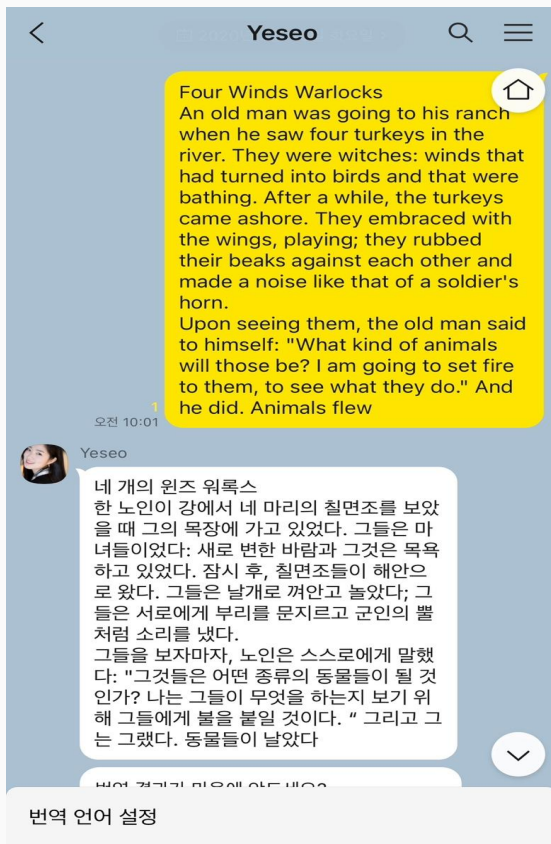
AI 번역 서버와 연계

유저가 입력한 메시지 번역

번역이 마음에 들지 않을 경우

집단 지성 번역 요청하여

더 나은 번역 결과를 받아볼 수 있음



카카오톡 챗봇

# 메신저 번역 챗봇

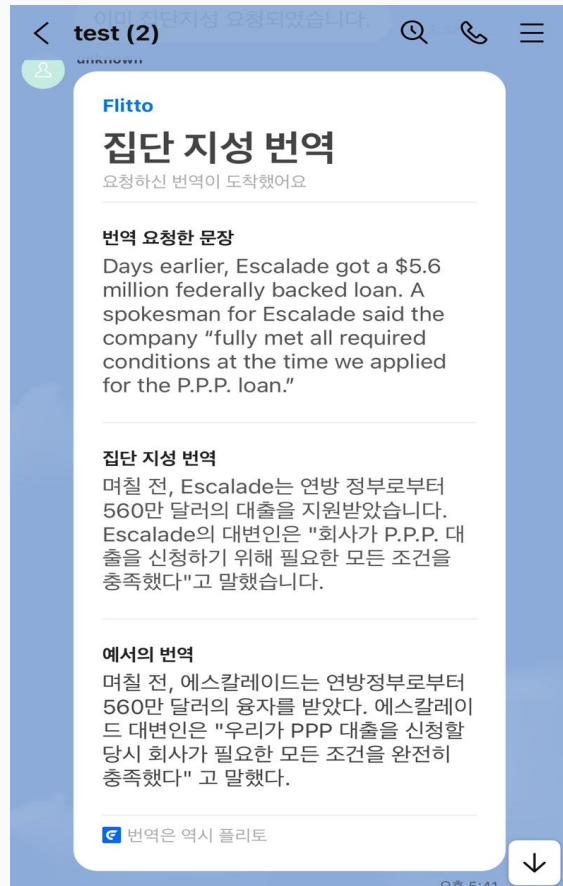
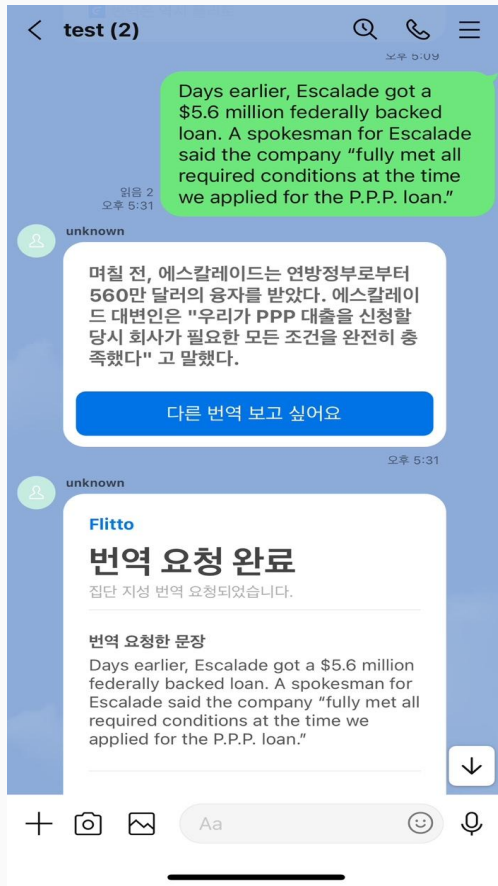
지원 메신저

라인, 카카오톡, 텔레그램, 슬랙, 페이스북

각 메신저 제공 Python API 사용

Django Server와 메시지 교환

사용 기술: Python, Django



라인 메신저 챗봇

# 텍스트 자동 분류 엔진

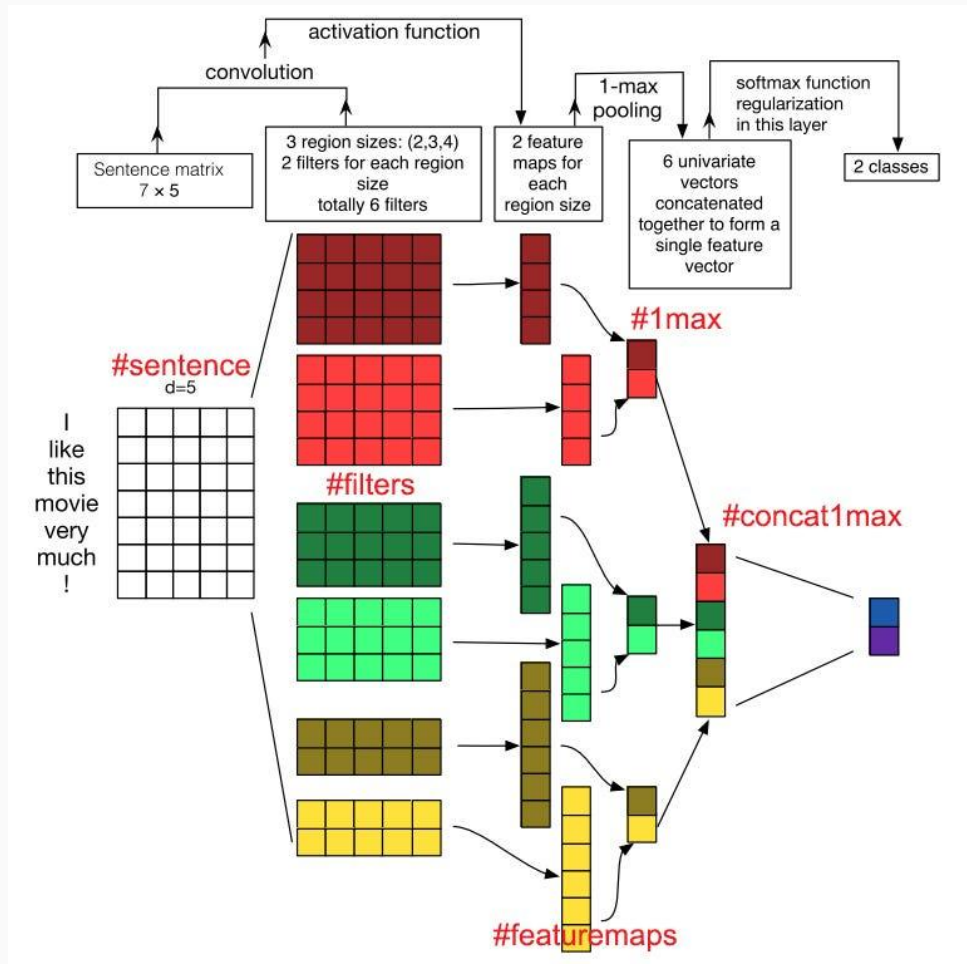
텍스트 데이터의 주제별 카테고리 분류

영어/한국어 분류 가능

최초 CNN 모델 구현 후 LSTM 모델 추가

그 후 BERT사용 모델로 고도화

사용 기술: Tensorflow Keras, Flask





# 음성 구간 감지 엔진

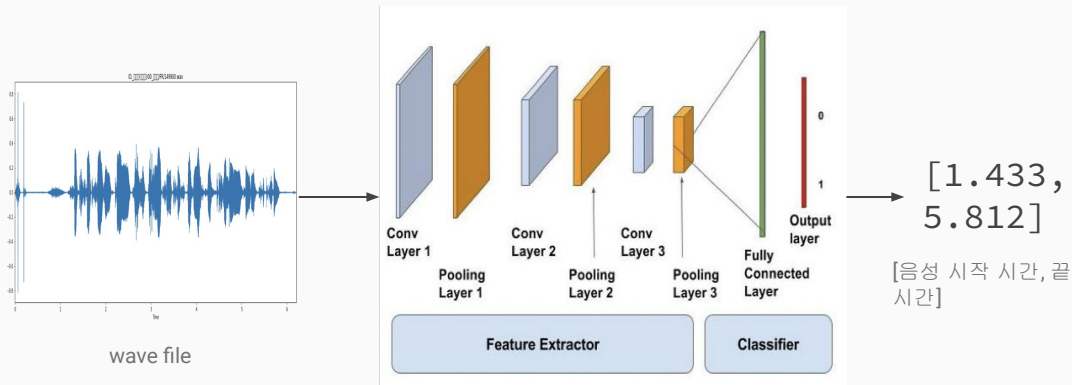
wav파일을 읽어서 음성 구간을 감지

음성 데이터 납품 전처리 용도

AWS Transcribe로 학습 데이터 생성

Keras CNN 모델 사용

사용 기술: Tensorflow Keras, Flask



# 언어 자동 감지 엔진

텍스트 데이터의 언어를 인식

AI 자동 번역의 언어 감지

납품 전 잘못된 언어의 데이터 필터링

fasttext 모델 + 자체 후처리 로직

특정 언어 감지용 모델 별도 학습하여 추가  
(말레이시아어 / 인도네시아어 등)

