Unity에서 동작하는 STT와 LLM 실습

렐루게임즈 딥러닝 팀 송주환
joowhan.song@relugames.com
렐루게임즈 테크 팀 김예진
evaline@relugames.com

2024.10.22

© 2023 ReLU Games, Inc.

1. 강의 소개

강의 순서

내용	시간
강의 소개	15분
STT 이론 및 실습	30분
LLM 이론 및 실습	30분
쉬는 시간	15분
자유 실습	60분

© 2023 ReLU Games, Inc.

1. 강의 소개

강사 소개

경력 사항

ReLU Games
정규직 · 1년 4개월
대한민국

- **딥 러닝 개발자**
  2024년 4월 - 현재 · 7개월
- **딥 러닝 팀 팀장**
  2023년 7월 - 2024년 4월 · 10개월
  대면재택혼합근무

  Deep Learning, Data Visualization 및 +2 보유기술

Deep Learning Engineer
KRAFTON Inc. · 정규직
2021년 5월 - 2023년 8월 · 2년 4개월
대한민국 서울

Robot Intelligence Engineer
TOROOC · 정규직
2019년 9월 - 2021년 5월 · 1년 9개월
Seoul, Korea

Data Scientist
OnePredict
2016년 10월 - 2019년 3월 · 2년 6개월
Seoul, South Korea

© 2023 ReLU Games, Inc.

1. 강의 소개

강의 목표

2개의 샘플 Scene으로 배워보는 Unity에서 동작하는 STT와 LLM

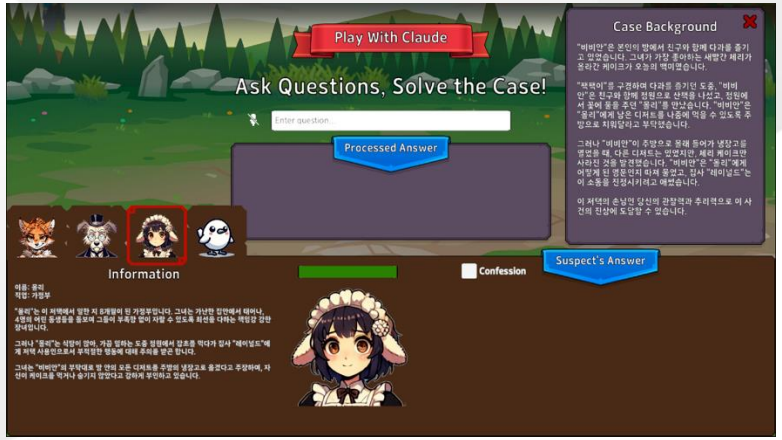
샘플 Scene과 코드를, 추후 수강생 각자 프로젝트의 밑바탕으로 삼을 수 있다

STT

Unity에서 음성 인식 기능을 구현한다

Unity Sentis를 이용해 Whisper 음성 인식 모델을 사용한다

Azure Language Studio API를 이용해 음성 인식 기능을 사용한다



LLM

Claude API를 사용한다

Claude API의 Prompting을 활용해본다 (json)

© 2023 ReLU Games, Inc.

1. 강의 소개

강의를 시작하기 전에

1. git clone github.com/Joovvhan/GGA2024ReLU
2. Unity로 프로젝트 열기
3. Assets/Scenes/STT Scene 열기
4. GGA2024ReLU 저장소의 구글 드라이브 링크에서 AudioDecoder_Tiny.sentis, AudioEncoder_Tiny.sentis 다운로드
5. Assets/StreamingAssets 폴더 내부에 다운로드한 파일 배치
6. STT Scene 정상 동작 확인 (Whisper Tiny, Azure STT)

7. Assets/Scenes/LLM Scene 열기
8. GGA2024ReLU 저장소의 구글 드라이브 링크에서 keys.txt 파일에서 key 확인
9. Assets/Scenes/LLM Scene의, ClaudeClient 오브젝트의 ClaudeClient 스크립트에 API_KEY 변수 입력
10. LLM Scene 정상 동작 확인 (Claude, Azure STT)

© 2023 ReLU Games, Inc.

2. STT 이론 및 실습

STT(Speech-to-Text)

음성 인식 모델 = 음성 데이터를 텍스트로 바꿔주는 모델

음성 데이터 = 일정한 시간 간격으로 녹음 장치에서 측정된 진폭의 시계열 데이터

텍스트 = 단어 토큰 (Assets₩StreamingAssets/vocab.json)

음성 데이터에서 텍스트로 변환: 가변 길이를 가변 길이로 변환 (Seq2seq)

© 2023 ReLU Games, Inc.

2. STT 이론 및 실습

Whisper

OpenAI에서 공개한 STT 모델
학습된 모델이 공개되어 있음
상대적으로 빠르고 정확함
한국어 지원

OpenAI Hugging Face

© 2023 ReLU Games, Inc.

2. STT 이론 및 실습

Hugging Face

기계 학습 관련 모델과 데이터를 공유하는 커뮤니티 = 모델과 데이터를 위한 Github

Unity Hugging Face

ONNX: 다양한 형식의 머신 러닝 모델의 공용 포맷

ONNX Runtime: ONNX 모델을 실행하는데 최적화된 패키지. C#에서 동작하나 Unity에 최적화 되어 있지 않음.

ONNX 시각화 도구: Netron

© 2023 ReLU Games, Inc.

2. STT 이론 및 실습

Unity Sentis

ONNX 파일을 Unity에서 돌리는데 최적화된 Unity 패키지

기존 Unity Barracuda를 대체함

Unity 6 이전 버전은 Package Manager에서 Sentis 설치 어려움

버전 요구 사항에 주의

Sentis 설치

© 2023 ReLU Games, Inc.

2. STT 이론 및 실습

Azure Speech Studio

STT, TTS, 음성 번역, 영상 자막, 녹음 등 MS에서 제공하는 음성 서비스를 이용할 수 있는 Azure 서비스

계정 만들고 무료로 이용하기

Speech SDK for Unity – 공식 샘플을 제공

Azure STT와 Whisper STT 비교

© 2023 ReLU Games, Inc.

2. STT 이론 및 실습

STT의 부정확성을 극복하는 방법

더 크고 정확한 모델을 사용한다

딥 러닝의 특징 – 최대 우도 (Maximum Likelihood)

한글 완성형(Hangul Syllables)과 한글 조합형(Hangul Jamo)

시퀀스와 시퀀스를 비교하는 방법 (Levenshtein Distance)

© 2023 ReLU Games, Inc.

2. STT 이론 및 실습

실습

자유롭게 코드를 수정하거나, 테스트해보세요

Whisper: WhisperEngine/RunWhisper.cs

Azure: AzureSTT/AzureSTT.cs

STT 결과물: DecisionSystem/DecisionSystem.cs (private void OnTranscriptionComplete)

Azure 계정을 만들고 무료 Key를 발급해봅니다

© 2023 ReLU Games, Inc.

3. LLM 이론 및 실습

LLM(Large Language Model)

언어 모델

텍스트 입력을 텍스트 출력으로 변환

입력과 출력 쌍이 모델의 목적을 정의함

번역 모델 / 대화 모델 (ChatGPT) / 요약 모델 / 요청 수행 모델 (InstructGPT)

상용 서비스: ChatGPT, Claude, Perplexity, Gemini, HyperCLOVA X

© 2023 ReLU Games, Inc.

3. LLM 이론 및 실습

Prompt Engineering

언어 모델에게 잘 부탁하는 법

[Anthropic Prompt Engineering]

[OpenAI Prompt Engineering]

ClaudeClient/ClaudeClient.cs

© 2023 ReLU Games, Inc.

3. LLM 이론 및 실습

정형화 된 출력 요청

json schema 요청

enum string 요청

조건부 실수 요청

boolean flag 요청

© 2023 ReLU Games, Inc.

3. LLM 이론 및 실습

입력 요청 – 토큰

토큰 = 돈 & 시간

Logs

Last refresh time: 2024년 10월 23일 오전 11:28 GMT+9

TIME (GMT+9)	MODEL	MODEL LATENCY	WORKSPACE	INPUT TOKENS	OUTPUT TOKENS	TYPE	ERROR
2024년 10월 22일 오후 4:17	claude-3-opus-20240229	5.40	Default	942	113	HTTP	
2024년 10월 22일 오후 4:17	claude-3-opus-20240229	3.05	Default	299	14	HTTP	

© 2023 ReLU Games, Inc.

3. LLM 이론 및 실습

실습

자유롭게 코드를 수정하거나, 테스트해보세요

Claude: ClaudeClient/ClaudeClient.cs

나만의 캐릭터를 만들어봅시다

토큰을 효율적으로 사용할 수 있는 방법을 적용해봅니다

Claude 계정을 만들고 Key를 발급해봅니다

© 2023 ReLU Games, Inc.

4. 자유 실습

실습 아이디어

Whisper (Tiny, Base, Medium)모델 성능 비교해보기

Whisper 음성 인식에 자동 녹음 종료 기능 반영

Claude 새로운 캐릭터성 구현

Claude 순차적 대화 해금 기능 구현

심화 실습 아이디어

ChatGPT API로 대체해보기

Azure TTS 기능 적용해보기

Sentis에서 Phi 1.5 사용해보기

© 2023 ReLU Games, Inc.

End of Document

© 2023 ReLU Games, Inc.