FINAL PROJECT - VIRTUAL MC

Group 3 - Applied Data Science

CONTENTS

1 Overview

2 File Organization

3 Functionality

CONTENTS

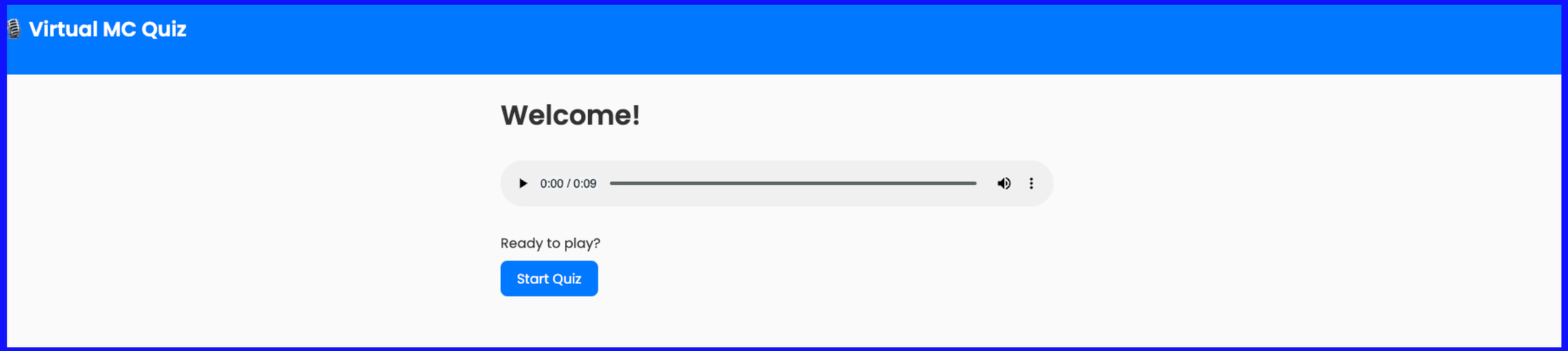
4 Models

5 Training Process

6 Demo

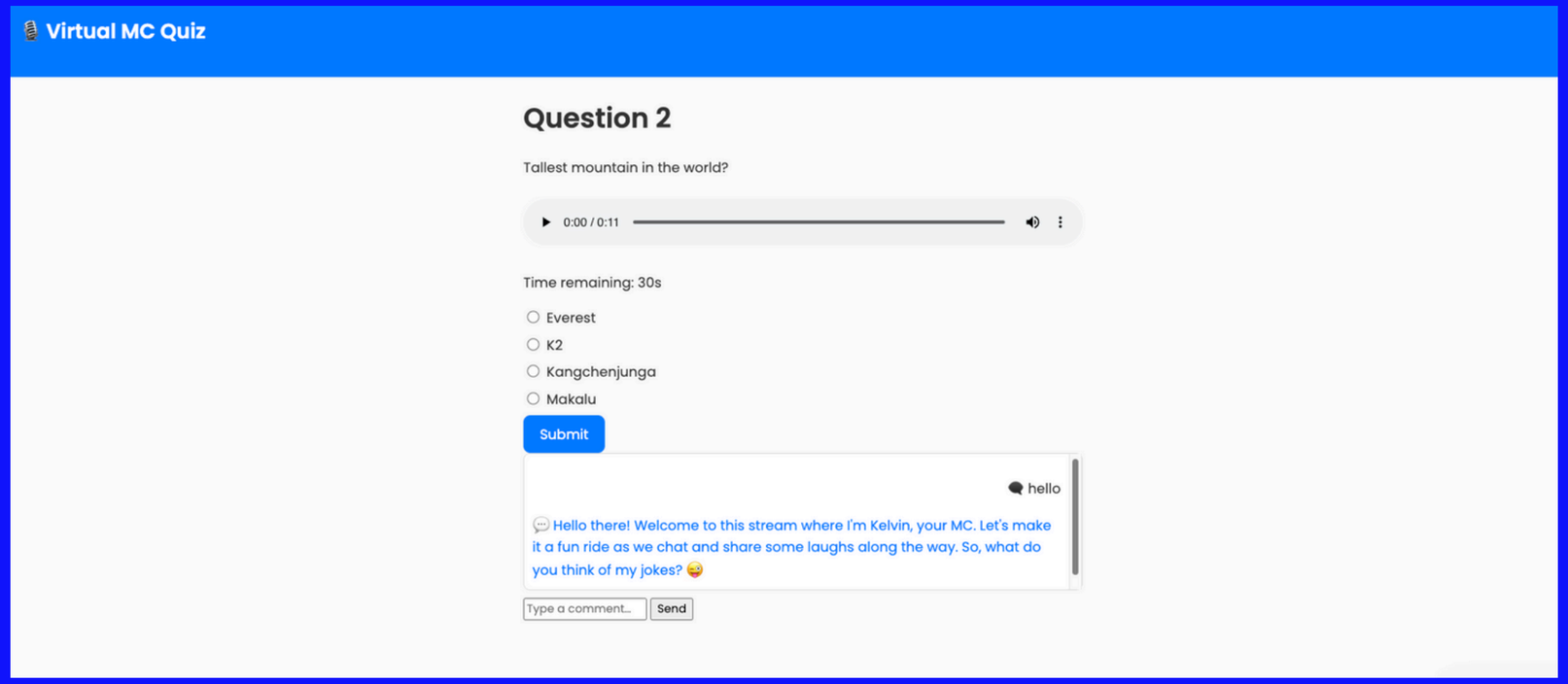
OVERVIEW

OVERVIEW



Starting screen

OVERVIEW



Quiz game screen

FILE ORGANIZATION

FILE ORGANIZATION

```
virtual_mc2/
│
├── llm/
│   ├── __pycache__/
│   ├── __init__.py
│   └── lm_client.py
│
├── static/
│   ├── audio/
│   ├── answer.js
│   └── quiz.js
│
├── templates/
│   ├── answer_feedback.html
│   ├── base.html
│   ├── index.html
│   └── quiz.html
│
├── tts/
│   ├── __pycache__/
│   └── local_f5/
│   ├── check.py
│   ├── config.yaml
│   ├── generate_audio.py
│   ├── output.wav
│   ├── server.py
├── main.py
└── requirements.txt
```

FILE ORGANIZATION

User Interface: HTML + CSS + Javascript

Deploying F5-TTS: FastAPI

Deploying LLM Model: LM Studio

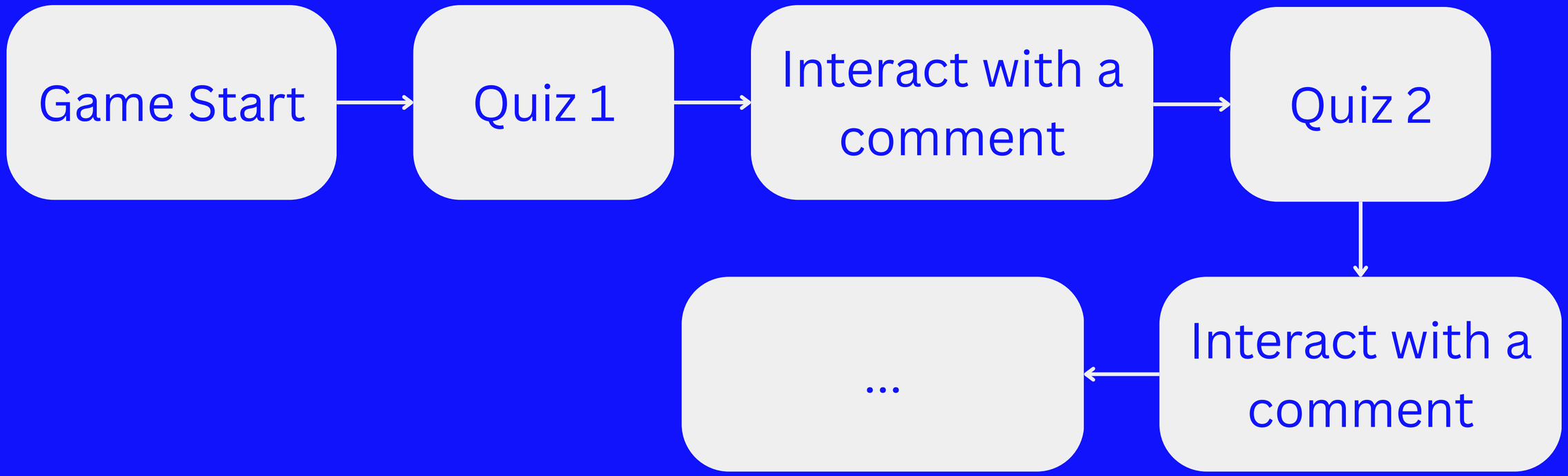
FUNCTIONALITY

FUNCTIONALITY

Quiz Game

Interacting with comments

FLOW


Game Start
Quiz 1
Interact with a comment
Quiz 2
Interact with a comment
...

MODELS

MODELS

F5-TTS

Nyagpt Qwen 2 7B v0.1

F5-TTS

model.safetensors → f5-tts-mlx → Run

f5-tts-mlx

trainer.py

```python
def load_checkpoint(self, step):
    params = mx.load(f"/Users/kelvinnguyen/Projects/virtual_mc2/tts/local_f5/model.safetensors")
    self.model.load_weights(list(params.items()))
    self.model.eval()
```

NYAGPT QWEN 2 7B V0.1

Particular for streaming answer in English

Optimized with prompting

TRAINING PROCESS

F5-TTS

Dataset: Emilia-English

Small sample : ≈25 h audio from 46k hours

Training period: ~17 days

F5-TTS

E2/F5 TTS Automatic Finetune

This is a local web UI for F5 TTS with advanced batch processing support. This app supports the following TTS models:

- F5-TTS (A Fairytaler that Fakes Fluent and Faithful Speech with Flow Matching)
- E2 TTS (Embarrassingly Easy Fully Non-Autoregressive Zero-Shot TTS)

The checkpoints support English and Chinese.

For tutorial and updates check here (https://github.com/SWivid/F5-TTS/discussions/143)

Tokenizer Type

Error

Project Name

my_speak

Create a New Project

Project

my_speak_pinyin

Refresh

Transcribe Data | Vocab Check | Prepare Data | Train Data | Test Model | Reduce Checkpoint | System Info

```
Skip this step if you have your dataset, metadata.csv, and a folder wavs with all the audio files.
```

Audio from Path

Voice

Drop File Here

- or -

Click to Upload

DEMO

DEMO VIDEO

Link

THANK YOU

FOR WATCHING