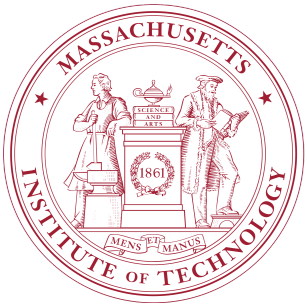
15.S04: Hands-on Deep Learning
Spring 2023



Project Report: JourneyJuice! An AI-based travel companion

Submitted by Team A.1: Michael Jiang (jjr310@mit.edu), Duanchen Liu (liudc@mit.edu), Anant Vashistha(anant101@mit.edu), Bingyu Zhou (bingyuz2@mit.edu)

May 15, 2023

Project Report: JourneyJuice! An AI based Travel Companion

15.S04 Hands-on Deep Learning

Submitted by Team A.1: Michael Jiang (jjr310@mit.edu), Duanchen Liu (liudc@mit.edu), Anant Vashistha(anant101@mit.edu), Bingyu Zhou (bingyuz2@mit.edu)

---

1. Description

Tourists frequently encounter a feeling of FOMO (fear of missing out) when they are precluded from hiring a tour guide due to monetary or linguistic obstacles, resulting in a dearth of pertinent information regarding the destination and the numerous events available to them. In view of this, there is a compelling necessity for a product that can equip travelers with the requisite knowledge and autonomy to comprehend a location's historical and cultural importance while also pursuing the prevailing events of interest in the vicinity. So, we are proposing a travel companion that can provide unforgettable experiences to travelers.

2. Objective

The project aims to build an AI product (travel companion) that can enhance travelers' overall experience visiting renowned tourist destinations by providing comprehensive information about the location and notifying them about the events happening in their vicinity.

3. Data Description

The AI product "JourneyJuice" deals with audio and textual data. The product uses the user's real-time data and scraps the information from the publicly available data through API calls.

1. Audio Data: Voice of the user

The input data is a human voice of fewer than 5 seconds (for example, "I'm in front of the Charles River"), as required to be handled by Whisper API to transform into text. Such voice data will not be part of the training data (Whisper and ChatGPT are already well-trained models).

2. Text Data: Events Data collected through web-scraping

The event information and their locations come from the Google Event page by calling google events API (using SerpAPI). This gives us information about events near an input location, including event name, description, place, etc.

4. Approach: Methodology & Flow of Idea

Our project aims to create a pipeline to process a traveler's audio input describing their place of visit and provide them with helpful information about the significance of the place, including nearby events in the traveler's preferred language. The events have been further categorized into six categories using the BERT model and thus give granular information.

The deep-learning pipeline will work as follows:

Step 1: Audio Input: The user provides an audio description of their location.

Step 2: Audio to Text Conversion: Use the Whisper API to transcribe the audio to text.

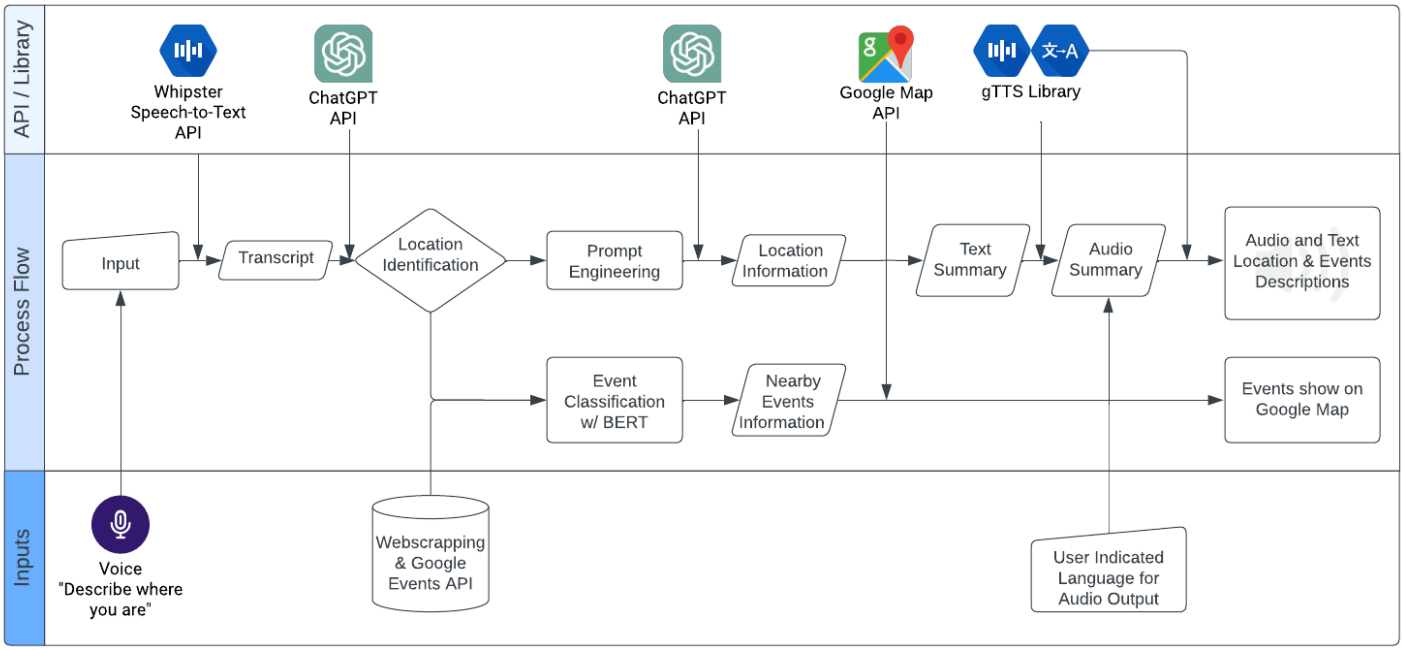
Step 3: Location Identification: Leverage the ChatGPT API to identify the location from the text transcript.

*Step 4:*Location Information: Use ChatGPT API to provide the significance and historical relevance of the place in fewer than 50 words.

Step 5: Events recommendation: Use web scraping techniques to gather information about same-day and same-week events in the area.

Step 6: Events classification: Use transfer learning by leveraging the pre-trained BERT model to classify the events based on event description in different categories.
Step 7: Text to audio: Present the information in a concise manner on the web application, with the option to convert text to audio.

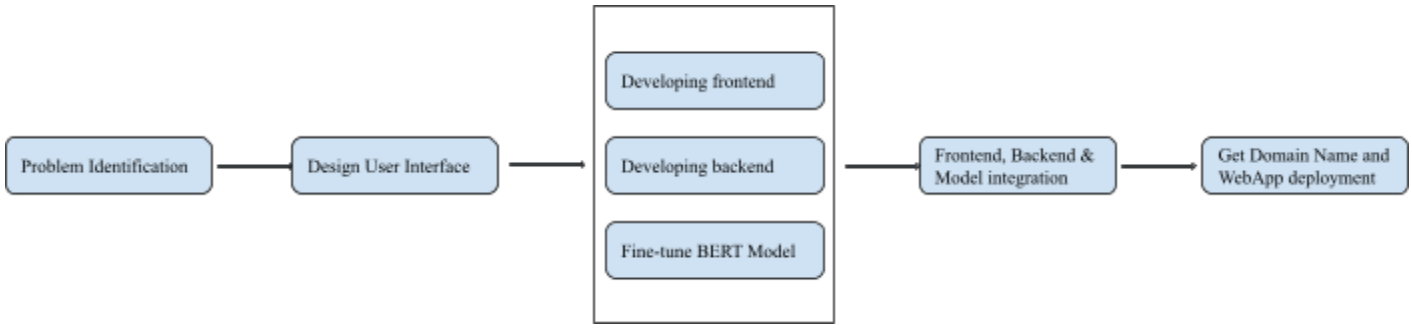
The detailed project process flows with location information on the input data, including the APIs and libraries leveraged given below.



5. Product Development Journey

The travel companion “JourneyJuice” is developed in the following five phases:
Phase 1: Problem identification
Phase 2: Design user interface
Phase 3: Develop frontend, backend, and BERT model
Phase 4: Frontend, Backend, and Model Integration
Phase 5: Get a domain name and deploy WebApp

The schematic view of the product development roadmap is shown below:



6. Deep Learning Model

6.1 Input Data

The team collected the event information and their locations from the Google Event page using events API. Then, the team labeled the events into six categories: Performance, Kids Event, Conference & Talk, Festival / Religious, Food & Drinks, and Sports. The team used 94 fine-tuning points and 20 testing points. (Sample data can be found in the Appendix)

6.2 Model

The team leveraged the power of transfer learning by using a pre-trained BERT model already trained on a massive amount of text data. The pre-trained BERT model has learned to recognize and capture the semantic meaning of words and sentences, which makes it an ideal starting point for a text classification task like event categorization.

To use the pre-trained BERT model for this task, the team fine-tuned it on the event dataset. Fine-tuning involves training the BERT model on a smaller, labeled dataset (in this case, the team's event data) to adjust its parameters for the specific task with Dense and Dropout layers. During fine-tuning, the model is exposed to the labeled data, and its weights are updated to learn how to classify events based on their descriptions.

The team used the fine-tuned BERT model to classify events into six categories. To do this, the team fed the description and title of each event to the BERT model, and the model output a probability distribution over the six categories, indicating the likelihood of the event belonging to each category. The team assigned each event to the category with the highest probability.

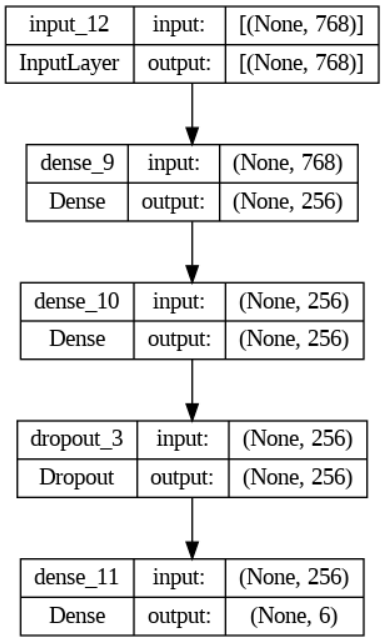
6.3 Model Architecture

To fine-tune a BERT model for a text classification task, the team used the pre-trained BERT model as a feature extractor and added a few layers on top of it for classification. The architecture of the BERT model for fine-tuning can be broken down into the following steps:

Input Encoding: The input text is first tokenized using the BERT tokenizer, which converts the text into a sequence of tokens. The tokens are mapped to their corresponding token embeddings using the BERT embedding layer.

Pre-Trained Encoder: The pre-trained BERT model is used as an encoder, which processes the token embeddings to produce contextualized word embeddings. The contextualized word embeddings capture the meaning of each word in the context of the sentence.

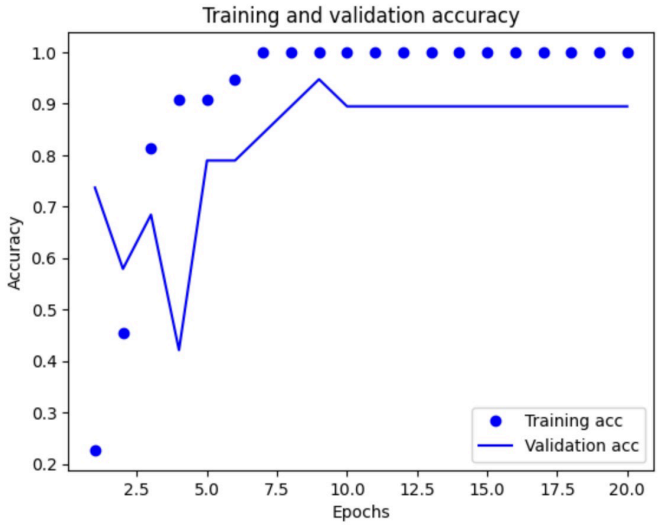
Classification Layer: A classification layer is added on top of the BERT encoder, which takes the contextualized word embeddings as input and produces the final output, the six categories of local events.



Fine-Tuning: The entire model is then fine-tuned using the events description data. The layers that the team added in the process are the Dense layer and the Dropout layer, with a total of 264,198 parameters. The weights of the pre-trained BERT model are frozen during the initial stages of training, and only the weights of the classification layer are updated. As the training progresses, the weights of the entire model are updated to fine-tune the model for the classification task.

In conclusion, BERT is a robust deep-learning architecture for NLP tasks, and fine-tuning BERT can lead to state-of-the-art performance. The architecture of the BERT model for fine-tuning involves using the pre-trained BERT model as a feature extractor and fine-tuning the model on the event classification task.

6.4 Model Results

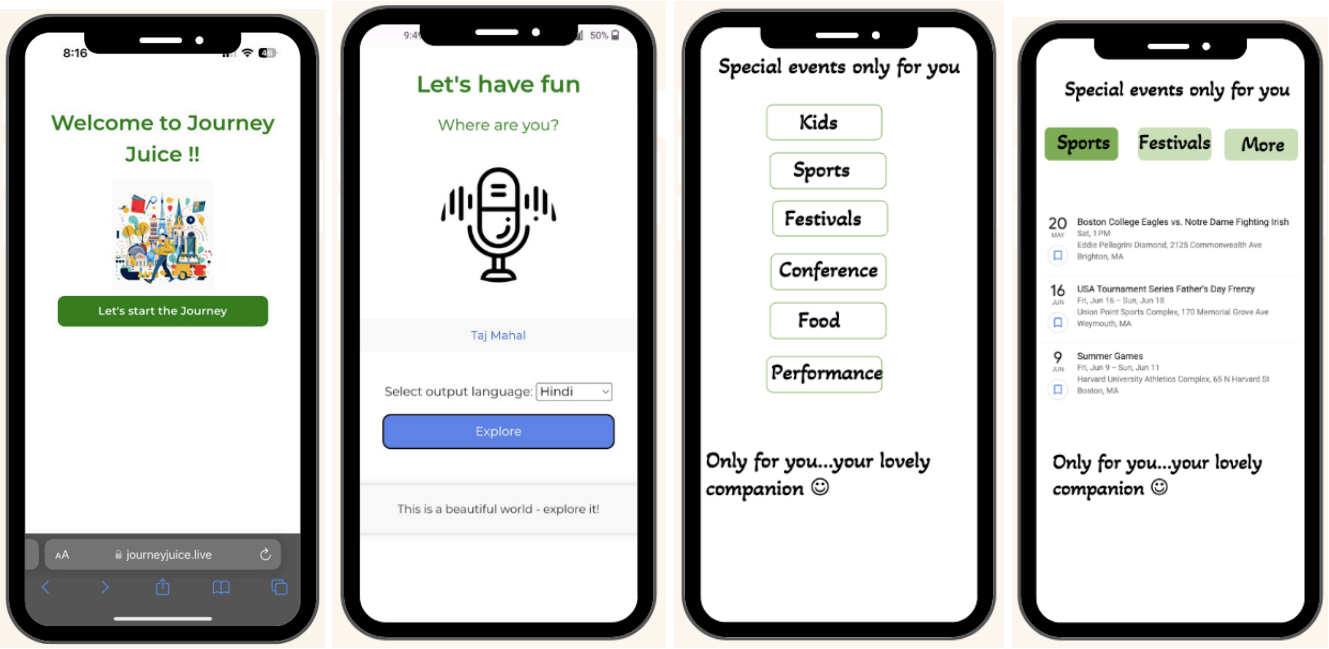


Overall, using a pre-trained BERT model for event classification can be a powerful and efficient approach, as it leverages the knowledge and expertise that has already been encoded in the pre-trained model.

The team used a 0.2 validation split, and the training and validating accuracy for each epoch are presented in the graph below. Fine-tuning the model on a smaller labeled dataset can quickly adapt it to the specific task at hand and achieve **0.95** testing accuracy in event classification.

7. Results & Impacts

The result of this project is a web application that can enhance travelers' experiences by providing them with comprehensive information about their destination in their preferred language and notifying them of nearby events. Please find the video demo for our current minimal viable product (MVP) here: https://youtu.be/2uhzuC2kSNk.



The product has a significant impact on traveler's life due to the following reasons.

JourneyJuice will enable travelers to explore new destinations and important events nearby, providing them with a more personalized and immersive travel experience. By leveraging deep learning techniques, the product can recommend events and activities based on the traveler's interests, preferences, and current location. This can help travelers discover new and exciting destinations and events they may attend.

Secondly, JourneyJuice can help promote local events and boost the local economy by attracting tourist attendance and promoting local community involvement. By recommending local events and activities, the product can help increase tourist attendance, which can positively impact local businesses, such as restaurants, shops, and accommodations. Additionally, by promoting local events, the product can help foster community involvement and engagement, which can positively impact the local culture and economy.

Thirdly, the product can also positively impact tourism and cultural exchange by bridging the linguistic and cultural gaps potentially hindering travelers' experiences. The product can help travelers better understand and appreciate the local culture, history, and traditions by providing real-time language translation and cultural context. This can also positively impact global relations and diplomacy.

Fourthly, the product can reduce dependence on tourist guides and eliminate language barriers. The product can help travelers navigate new destinations and participate in local events without needing a tourist guide or language interpreter by providing real-time language translation and event recommendations. This can help travelers feel more independent and empowered during their travels, which can positively impact their overall travel experience.

Finally, the product can help achieve better planning of events and logistics. The product can help travelers plan their trips more efficiently and effectively by providing real-time event recommendations and logistics information. This reduces stress and frustration associated with travel planning and logistics, which can positively impact travelers' overall satisfaction with their travel experience.

Overall, JourneyJuice has the potential to have a significant impact on travel and tourism by providing personalized recommendations, promoting local events and culture, bridging linguistic and cultural gaps, reducing dependence on tourist guides, and improving travel planning and logistics.

8. Lessons Learned: Challenges & Learnings

8.1 Challenges

The team encountered several challenges in this project.

Audio to Text Conversion: The first challenge was related to capturing a user's audio and converting it into accurate text. As the system relied on capturing the traveler's voice accurately, any errors in the voice recognition could lead to misspellings and inaccurate information. The team explored different voice recognition models to address this challenge to improve accuracy. They may have also explored error correction techniques. This issue was coming up due to accent bias, and the team explored the possibility of leveraging a spellchecker.

Limited availability of events data: The second challenge was related to the limited availability of event data. This limited the training and testing process of the BERT model in the events categorization process. The model may be further fine-tuned on the availability of datasets.

Heavy reliability on Google events: To provide real-time and accurate event recommendations to travelers, the system needed to rely on a reliable and up-to-date source of event data. In this case, the team relied on Google's event data, which may have had limitations in terms of the number and types of events available. The algorithm should also ensure that the event should be within a limited time period of a user trip.

API limitations: Additionally, there may have been limitations on the number of API requests that could be made due to budget constraints. Overall, ensuring the reliability and accuracy of the event data was critical to providing timely and relevant information to the traveler, and the team carefully considered and addressed these challenges during the development process.

Product development: The team faced the challenge of developing the actual product in the integration phase. The BERT model integration, along with the frontend and backend (API calls), required much team effort and a lot of hit and trial. Also, the website deployment on the server and linking it to the domain name to make an actual MVP of a product came across many programming issues.

8.2 Learnings

There are two main key lessons that the team learned in this project. These ideas are around the AI application development and leveraging ML models in production.

BERT Model Fine Tuning:

One key lesson in this project was the importance of fine-tuning in creating effective deep-learning models for language-related tasks. In this project, the team used NLP techniques to process and analyze large amounts of text data. For fine-tuning, the team understood the importance of data collection and data quality for the model's performance.

The team needed to gather and process real-time data on current events and activities to develop an effective deep-learning model. The quality and relevance of the training data could significantly impact the accuracy and effectiveness of the model.

Product Development: Another important lesson learned in this project was the importance of effective product development. In order to create a successful product, it was essential to understand the needs and desires of the target audience and to design a user-friendly interface that met those needs. The team worked closely with users throughout the development process, soliciting feedback and iterating on the product design to ensure that it was user-friendly and met the target audience's needs. Additionally, the team recognized the importance of scalability in product development. They designed the system to be easily expandable and adaptable to new locations and languages, allowing for future growth and expansion. Also, the team understood the importance of partnerships and collaboration in product development. The team worked closely with each other to ensure that the content provided by the system was culturally sensitive and appropriate.

9. Appendix

The links to the important references are as follows:

1. GitHub
2. Project Proposal
3. Google Collab
4. Product Introduction
5. Product Demonstration
6. Project Slides
7. WebApp

The product can be accessed through WebApp and is currently deployed on a Microsoft server.

Sample Events Data

Title	Description	Category
Art Attack	Kids will unleash their creativity and imagination through a variety of art projects and activities. They will experiment with different mediums and techniques and create their own masterpieces.	Kids Event
Craft Beer Tasting at MIT	Sample a wide variety of craft beers from some of the best breweries in the region. Learn about the brewing process, taste unique flavor combinations, and enjoy a night out with fellow beer enthusiasts.	Food & Drinks
Encountering the Divine through Multiple Faiths	While we often associate Jesus with Christian traditions, how might we encounter Jesus (and/or the Divine more broadly) through faiths beyond Christianity?	Festival / Religious
Ethical AI Summit	As artificial intelligence (AI) continues to advance, this summit explores the ethical and societal implications of these technologies. Speakers will discuss topics such as bias in algorithms, privacy concerns, and the role of AI in healthcare.	Conference & Talk
Film Festival	A celebration of filmmaking featuring short films, documentaries, and animations created by MIT students. The festival includes screenings, Q&A sessions with filmmakers, and awards ceremonies.	Performance