

What The Face (Advanced)

Inspiration

At What The Face, we realize that not everyone experiences emotions the same way. In fact, many people have a hard time perceiving, empathizing with, or reciprocating emotions because of differences in cognition. Perceiving emotional information from conversation does not come naturally to many individuals with Autism Spectrum Disorder (ASD). This is coupled with health disparity within ASD; many individuals from lower socioeconomic status (e.g., ethnic minorities, low-income families, etc.) never receive adequate, or comparable (to those from less-disadvantaged backgrounds), primary care and/or mental health care. However, with the relatively high access to smartphones in the U.S., including in those from disadvantaged backgrounds, the mobile health (mHealth) technology movement shows promise to lessen health disparity. What The Face is an attempt to utilize artificial intelligence to potentially help those with ASD determine the emotional context of social situations. Our project picks up facial emotion recognition and analyzes polarity of speech (using speech-to-text coupled with sentiment analysis) to help those on the autism spectrum, or any who struggle with understanding emotional context, understand and learn emotions used in context.

What it does

What The Face prompts the user to take a picture and record audio on the app's UI. This image is then run through Google Cloud Vision API to analyze the facial emotions of the person in the picture, which categorizes different emotions and ranks them on a scale of very unlikely to very likely. The audio file is captured through pyaudio and

converted to a string through the Google Cloud Speech Recognition API. This string is then sent to TextBlob API, which uses NLTK (Natural Language ToolKit) to analyze the polarity and subjectivity of the string. All of this information is then relayed to the user via the app's user interface.

How we built it

We used JavaScript, Node JS, React Native, Android Studio Code Expo XDE for our front end. For our back end, we used Google Cloud Vision API, Google Cloud Speech Recognition API, TextBlob API (NLTK), pyaudio, portaudio19, and python-all-dev (python2). We also made a landing page for our app with Wix ([link](#)).

Challenges we ran into

Once we figured out what we wanted to do, we split up and started working on our front and back end. Overall, we had a lot of problems with proving authentication and getting the correct permissions from the right places. For front end, we had a lot of challenges with figuring out how open a camera for our UI and take a picture. We also had problems with saving the image after it was taken. Very few of us actually had ever worked with react native before, and it took a while to understand the technology. For back end, we had a lot of challenges getting the correct authentication for our Google APIs. We had to download multiple versions of python and pip on multiple different computers to figure out which versions worked with which APIs. After hours of trying to use Google Cloud's Sentiment Analysis API, we decided to go with TextBlob.

Unfortunately since we coded our front end with react native and our back end with

python, we had much difficulty trying to get the two to connect. The picture gets sent to the server and the server successfully receives it, but our back end could not find the exact location of where the data is stored.

Accomplishments that we're proud of

We are very proud of the individual parts that we got to successfully run. None of us had used a Google Cloud API before, so we are proud of successfully learning, implementing and utilizing Google Cloud's Vision and Speech API, TextBlob, and pyaudio. We are also happy with our UI, which can successfully open the phone's camera and take/save a picture. We are proud of the functionality that our app's individual parts have, and we are excited for the promising prospects that What The Face has to advance the learnings of those who do not understand social and emotional context.

What we learned

We all learned a variety of skills through this project. Everyone in our team had contrasting skill sets when we started, but we got to share and collaborate these skills to teach and learn each other. We also used our resources wisely and spent much of our time learning from mentors. These skills range from learning the React Native framework, learning how to implement APIs, and even learning a new language for some (Python). In the end, we learned how to work with our group dynamic to collaborate and learn from each other.

What's next for us

Although we got a lot of individual parts working, we still need to connect our front end with our back end code. We would like to successfully implement Google Cloud's Sentiment Analysis API, as we believe it can be more accurate and provides more metadata than TextBlob's API. We would also like to be able to stream videos straight into our back end instead of a single picture, as it could provide a better social context for the situation. We would also like to implement a response assistant, prompting users who are uncertain on how to respond to a situation.