**HOME CREDIT**

**Business Objective:**

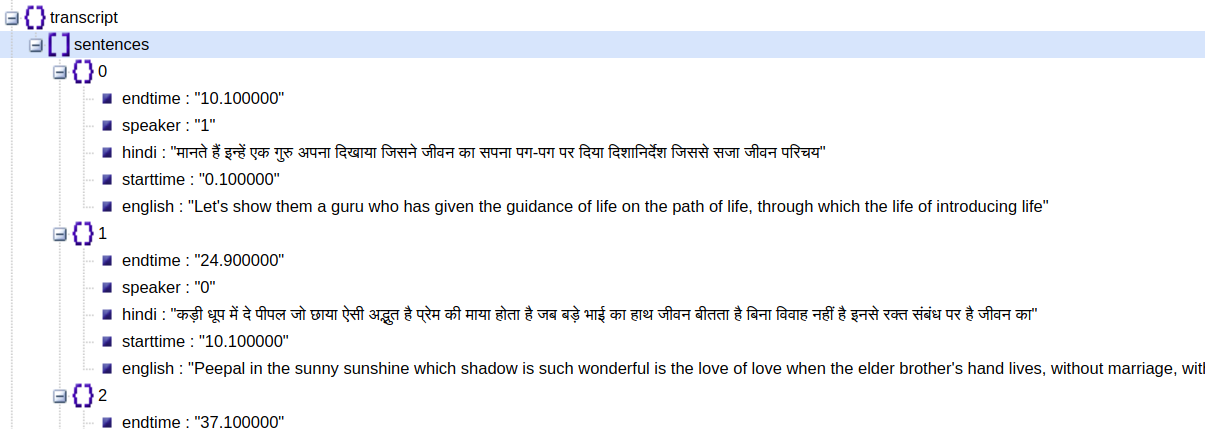
To convert customer service interaction recordings in Hindi to English transcripts for easier storage and analysis. The output format is a JSON consisting of sentences and speaker for each sentence.

**Tools & Technology Used** :   
**Backend:** Python 2.7 (PyAudioAnalysis, Flask), Google Cloud Speech API

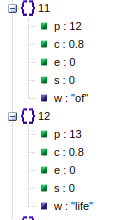
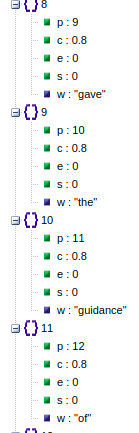
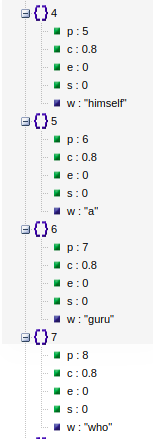
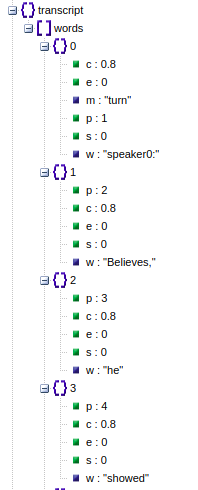
**UI:** HTML, CSS, Bootstrap, JavaScript

**Solution Approach:**

* The first step was speaker identification or speaker diarization. For this we used pyAudioAnalysis - a python library for a wide range of audio analysis tasks like feature extraction, classification and segmentation.
  + We used unsupervised segmentation (diarization) for speaker identification.
  + We converted all test recordings to .wav format which is suitable for use with pyAudioAnalysis. The files had two speakers like most customer care conversations.
  + This step generated segmented files of speakers in the order of their speaking.
* Google Cloud Speech API was used to convert the audio segments produced in the previous step to english text in JSON format. The segments are sent and the response is in the following format. “speaker” field specifies the speaker identification id. “starttime” & “endtime” are the start time & end time of the audio segment respectively. “hindi” field contains the sentence in Hindi converted from audio format by Google Speech API and “english” field is the translated text.



* We changed this format of output into a word wise output instead for ease of use by subsequent analysis tools and programs. A word-by-word format was created as shown below



* Each individual word has confidence tag(c), word number in the sentence(p), speaker turn tag(m) and translated word (w).
* Python package Flask was used to create a REST server which takes audio file names and path and provides a response JSON as shown above.

**Outcomes:**

* Diarized audio segments which can be used for verification or any form of audio analysis.
* JSON output of translated transcripts word-by-word.
* REST service for the translation system keeping ease of use in mind.

**Customer Benefits:**

* Fully automated transcript generation and translation system as a REST service
* The translation system can be integrated easily into any text analysis or customer service evaluation system.