Mohini Banerjee

Shilika Gehlot

Jessica Meng

iRecruit | Team B2

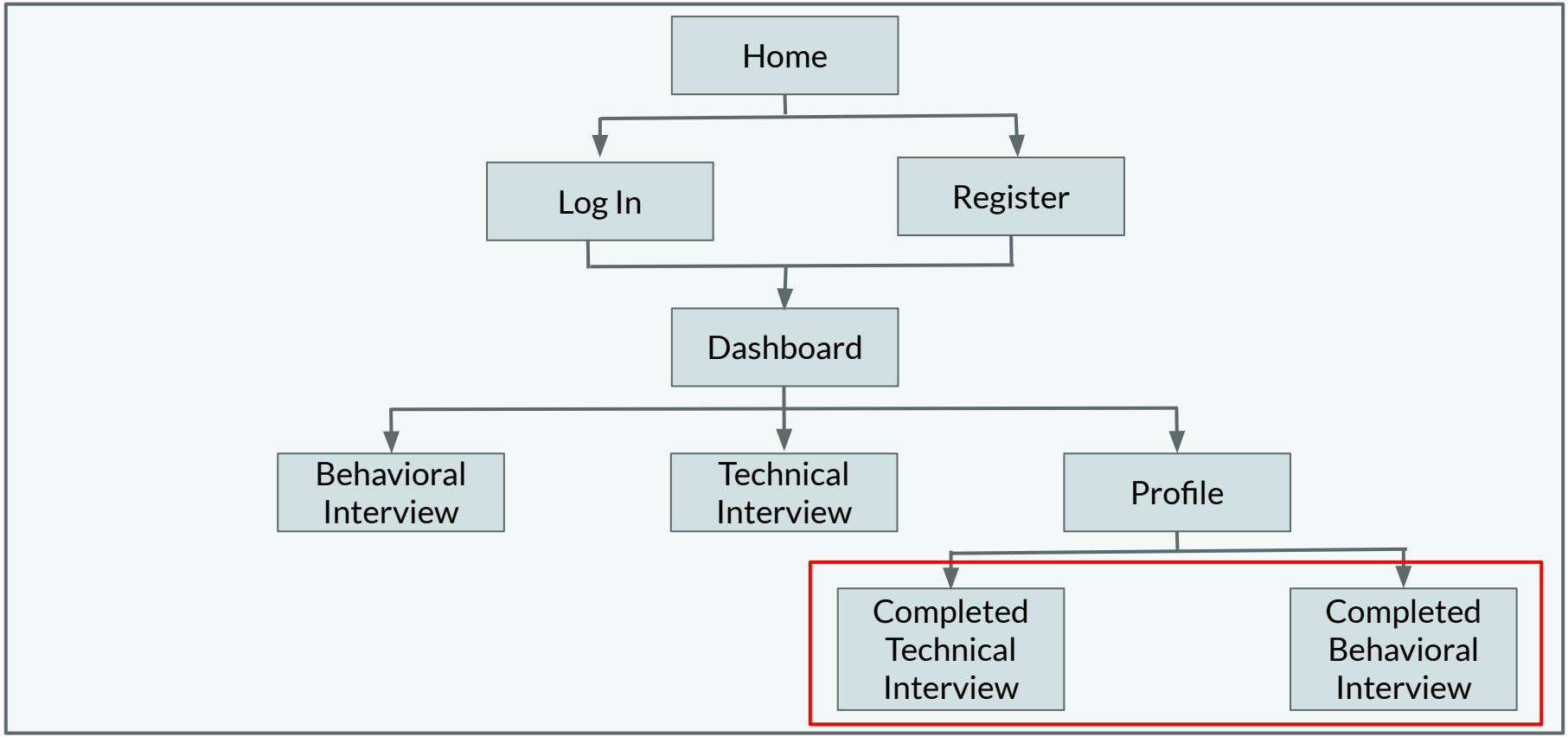
Application Area

- **Problem Area:** Lack of opportunity to practice a simulated, real-time interview
- **iRecruit:** Interview assistant capable of providing software engineering job-seekers raw interviewing experience
    - Behavioral and technical interview portions
- **Areas:** Software and Signals

Solution Approach

- **Behavioral:**
  - User video records themselves answering randomly generated questions
  - iRecruit provides real-time feedback for subpar eye contact and screen alignment
- **Technical:**
  - User audio records themselves saying question category (e.g. array, Java)
  - iRecruit builds speech recognition algorithm from scratch to determine category user is asking for
- User-friendly way to help prepare for interviews
- Centralized platform for both behavioral and technical portions

Complete Solution - Overall Web App



Complete Solution – Facial Detection

Web page with overview and instructions

Behavioral Interview Practice

iRecruit provides 3 options for behavioral interview practicing to account for various levels of experience. In all 3 options, users will video record themselves answering a randomly generated, common behavioral interview question. Please read further to determine which option is best for you. We recommend practicing against an empty background. The yellow circle in the center of the screen serves as a guideline for centering your face. We recommend trying to align the center of your nose to the circle. Begin answering the question when you see the red "RECORD" word appear at the bottom right corner of the screen.

Option 1 is to practice with both eye contact and screen alignment, in which iRecruit will alert you of subpar eye contact or screen alignment. We recommend practicing with this option for beginner-level users, who are either unfamiliar with the iRecruit behavioral interviewing platform or behavioral interviews in general.

Option 1: Eye Contact and Screen Alignment

On-click

Option 2 is to practice with only eye contact, in which iRecruit will alert you of subpar eye contact. We recommend practicing with this option for intermediate-level to advanced-level users, who are aware of their tendencies to move their eyes around a lot during behavioral interviews.

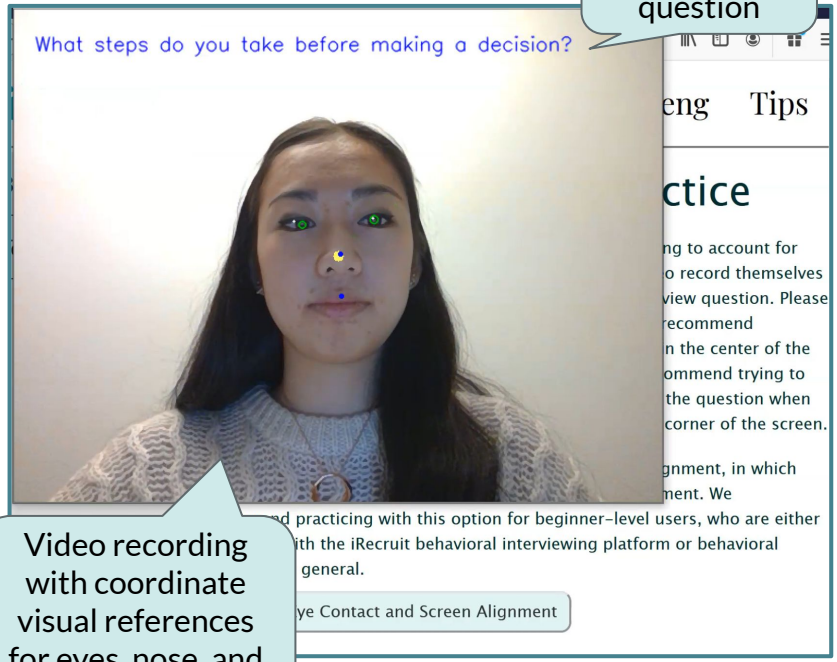
Option 2: Eye Contact Only

Option 3 is to practice with only screen alignment, in which iRecruit will alert you of subpar screen alignment. We recommend practicing with this option for intermediate-level to advanced-level users, who are aware of their tendencies to move their body off of the center of the screen during behavioral interviews.

Option 3: Screen Alignment Only

3 options to practice with to account for various levels of experience

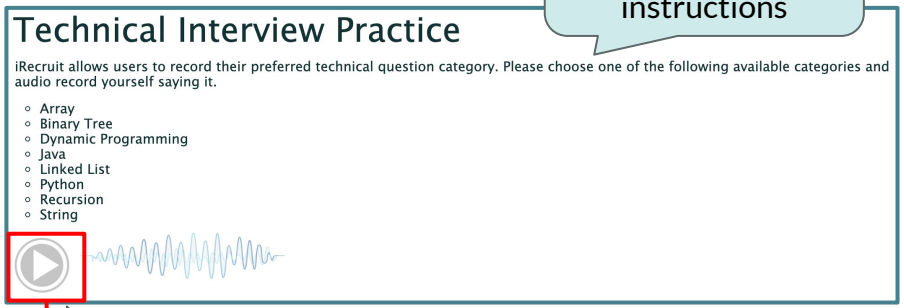
Randomly generated question



Video recording with coordinate visual references for eyes, nose, and mouth

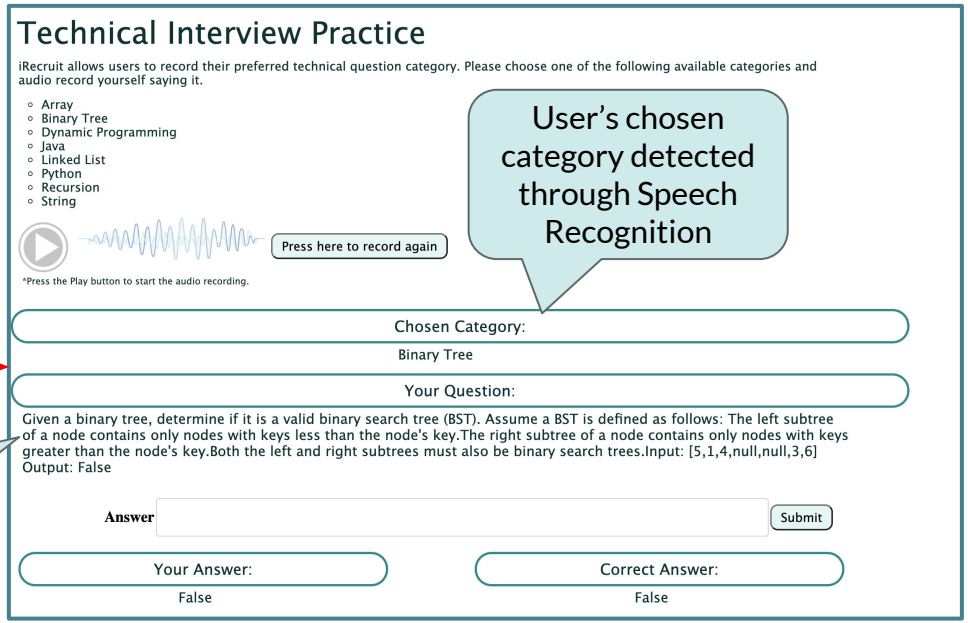
Complete Solution - Speech Recognition

Web page with overview and instructions



On-click

Voice recording button to record user's preferred technical skill



User's chosen category detected through Speech Recognition

Question generated based on chosen skill

Metrics Overall Accuracy Summary

Portion	Desired Accuracy	Actual Accuracy
Facial Detection Option #1 (Eye Contact and Screen Alignment)	80%	84.91%
Facial Detection Option #3 (Eye Contact Only)	80%	87.04%
Facial Detection Option #3 (Screen Alignment Only)	80%	98.04%
Speech Recognition	65%	35%

Design Tradeoffs

- **Facial Detection:**
  - OpenCV vs. dlib:
    - Chose OpenCV for experience level, Haar cascade accuracy, and documentation available
  - Behavioral Interviewing Techniques:
    - Forwent original plan of posture, no sufficient way to measure (no mouth == no face, shoulders may be covered by long hair)
- **Speech Recognition:**
  - Neural Network vs. Gaussian Mixture Model
    - Chose Neural Network for experience level and implementation
  - Letter-by-Letter Classification vs. Word Classification
    - Chose Word Classification for project scope and user practicality

Testing Approach

- Manual testing - keep track of actual vs. expected values
  - Facial Detection: # of times system alerts user of subpar eye contact and/or screen alignment based on predetermined set-up
    - Input: Facial feature coordinates
    - Output: Alert user of off-center coordinates within 5 seconds == passing test
  - Speech Recognition: # of words that system correctly recognizes
    - Input: Spectrogram representation of word
    - Output: Correct predicted word == passing test
- Accuracy measured by # of successful tests passed (true positives/negatives)

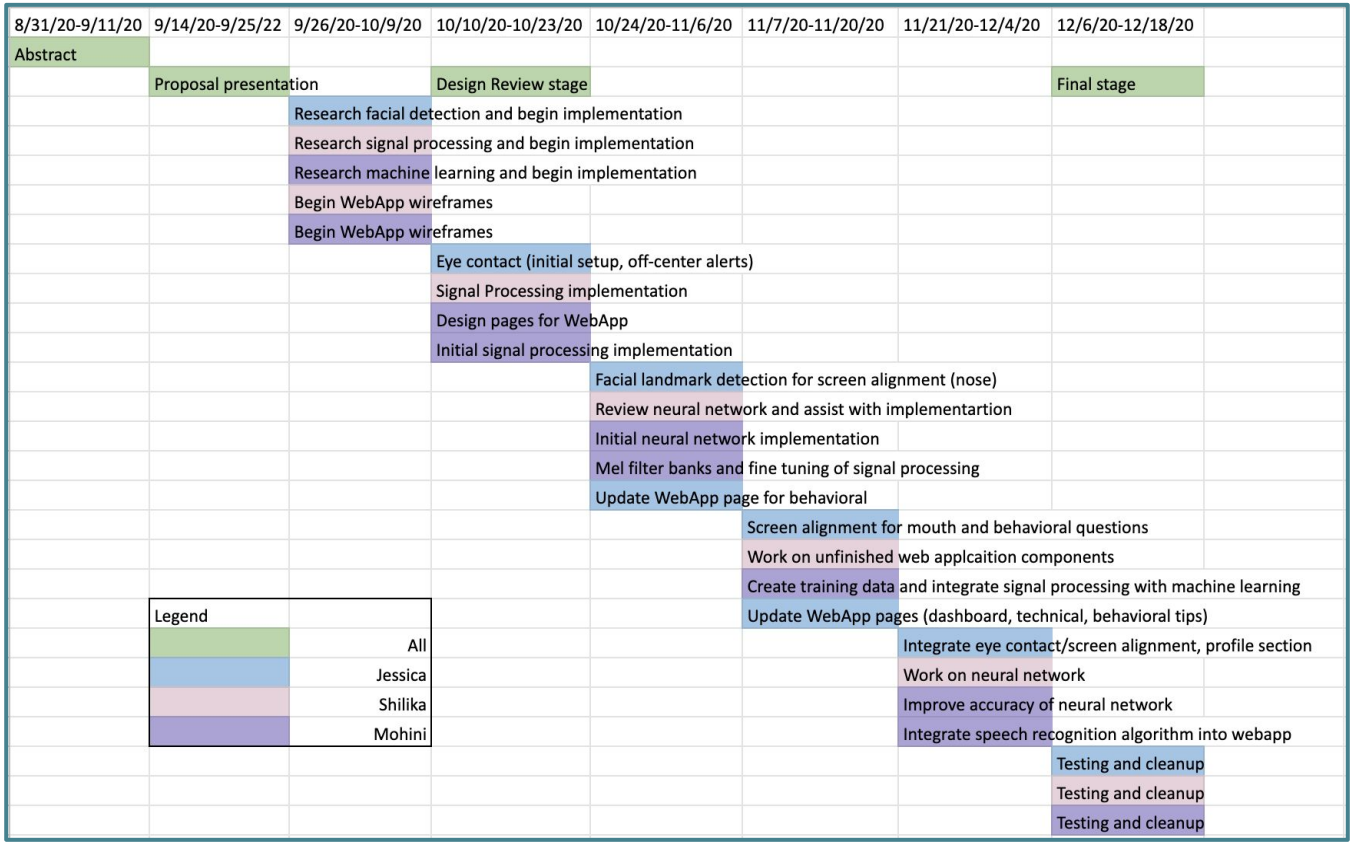
Facial Detection Metrics

Requirement	Description	Expected	Actual
Initial Set-Up	User is given ~5 sec to position themselves for us to learn facial features	<= 5 seconds	<= 5 seconds
Alerts	Alert user of subpar eye contact or screen alignment within 5 sec	<= 5 seconds	<= 5 seconds
Option #1 (Eye Contact and Screen Alignment) Accuracy	Measured by # of true positive/negative alerts of subpar eye contact and screen alignment	80%	84.91%
Option #2 (Eye Contact Only) Accuracy	Measured by # of true positive/negative alerts of subpar eye contact	80%	87.04%
Option #3 (Screen Alignment Only) Accuracy	Measured by # of true positive/negative alerts of subpar screen alignment	80%	98.04%

Speech Recognition Metrics

Requirement	Description	Expected	Achieved
Represent audio file as a waveform spectrogram	Use signal processing techniques to process audio recording	The same word spoken by the same person === same spectrogram representation	Yes
Neural Network Classification Accuracy	Build a neural net that accurately determines the probability distribution over the 8 output categories	65%	No

Schedule


8/31/20-9/11/20
9/14/20-9/25/22
9/26/20-10/9/20
10/10/20-10/23/20
10/24/20-11/6/20
11/7/20-11/20/20
11/21/20-12/4/20
12/6/20-12/18/20
Abstract
Proposal presentation
Design Review stage
Final stage
Research facial detection and begin implementation
Research signal processing and begin implementation
Research machine learning and begin implementation
Begin WebApp wireframes
Begin WebApp wireframes
Eye contact (initial setup, off-center alerts)
Signal Processing implementation
Design pages for WebApp
Initial signal processing implementation
Facial landmark detection for screen alignment (nose)
Review neural network and assist with implementartion
Initial neural network implementation
Mel filter banks and fine tuning of signal processing
Update WebApp page for behavioral
Screen alignment for mouth and behavioral questions
Work on unfinished web applcaition components
Create training data and integrate signal processing with machine learning
Update WebApp pages (dashboard, technical, behavioral tips)
Integrate eye contact/screen alignment, profile section
Work on neural network
Improve accuracy of neural network
Integrate speech recognition algorithm into webapp
Testing and cleanup
Testing and cleanup
Testing and cleanup
Legend
All
Jessica
Shilika
Mohini