

**CMPE 187 - Software Quality Engineering****Spring 2023 - Group Project****Deliverable 1 - Conventional Test Report**

Team member:

Khai Nguyen

Faith Deanon

Kevin Campbell

Long Nguyen

|  |           |
|--|-----------|
| <b>Section #1: Introduction</b>                          | <b>3</b>  |
| 1.1 App Information                                      | 3         |
| 1.2 Test Information                                     | 4         |
| 1.3 Task Partition                                       | 4         |
| 1.4 Project schedule                                     | 4         |
| <b>Section #2: Test requirements</b>                     | <b>5</b>  |
| AI function requirements                                 | 5         |
| <b>Section #3 – Selected test models and methods</b>     | <b>6</b>  |
| 3.1 Test methods   | 6         |
| 3.2 Decision table testing                               | 6         |
| 3.3 Equivalence partitioning                             | 7         |
| 3.4 Scenario testing                                     | 7         |
| <b>Section #4 – Test case design with test data</b>      | <b>7</b>  |
| 4.1 Decision table testing                               | 7         |
| 4.1.1 Detailed Decision Table                            | 8         |
| 4.1.2 Decision Table Test Cases                          | 10        |
| 4.1.3 Decision Table Test Coverage                       | 18        |
| 4.2 Equivalence testing                                  | 18        |
| 4.2.2 Equivalence Partitioning Test Cases                | 19        |
| 4.2.3 Equivalence Partitioning Test Coverage             | 28        |
| 4.3 Scenario testing                                     | 29        |
| 4.3.1 Scenario Testing Designs                           | 29        |
| 4.3.2 Scenario Testing Test Cases                        | 32        |
| 4.3.3 Scenario Testing Coverage                          | 39        |
| <b>Section #5 – Test result analysis and bug summary</b> | <b>39</b> |
| 5.1 Test Results Summary                                 | 39        |
| Decision Table Test Summary                              | 39        |
| Equivalence Partitioning Test Summary                    | 40        |
| Scenario testing test summary                            | 41        |
| 5.2 Manual Test Costs                                    | 41        |
| 5.3 Complexity and coverage                              | 41        |
| 5.4 Bug analysis and summary                             | 41        |
| Bugs   | 41        |
| Detecting feelings                                       | 42        |

# Section #1: Introduction

## 1.1 App Information

Wysa is an AI-powered chatting app to help users with anxiety by providing mental health support. In addition to free support through chat, it also offers exercises to reduce anxiety, paid coaching plans with actual experts, and self-care exercises.

However, the testing will only cover AI software chatting capabilities, especially how it can handle mental health conversations.

## 1.2 Test Information

We apply different types of black box testing in order to evaluate the effectiveness of Wysa's AI.

## 1.3 Task Partition

| Member name    | Task  |
|----------------|---|
| Khai Nguyen    | Decision table testing, Scenario Testing        |
| Faith Deanon   | Decision table testing, Scenario Testing        |
| Kevin Campbell | Equivalence partition testing, Scenario Testing |
| Long Nguyen    | Equivalence partition testing, Scenario Testing |

## 1.4 Project schedule

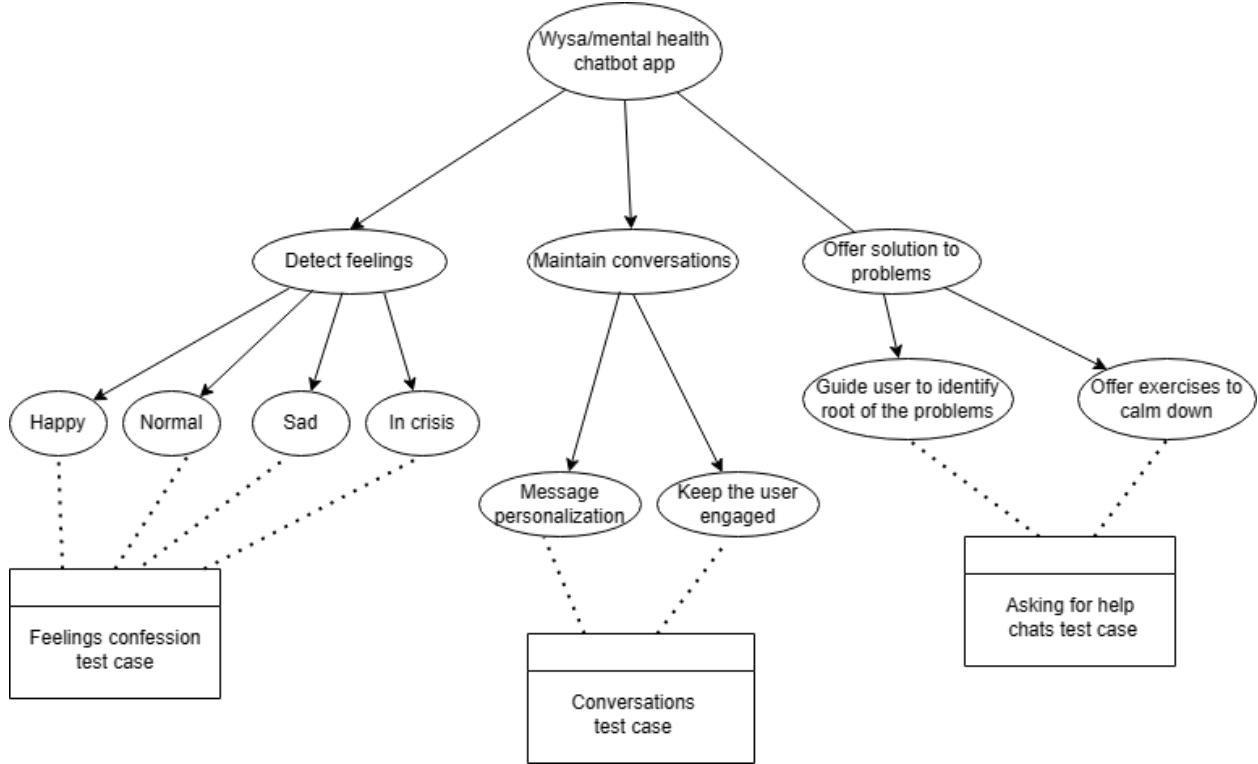
| Dates           | Description             |
|-----------------|-------------------------|
| Mar 1 - Mar 6   | Research about the app  |
| Mar 7 - Mar 10  | Come up with test cases |
| Mar 11 - Mar 16 | Test and collect data   |
| Mar 17          | Submission              |

## Section #2: Test requirements

### AI function requirements

The app is tested on both Android and iOS platforms.

For the chatting function, below is the AI requirements tree. The tests are designed to test the features listed on the tree.



## Section #3 – Selected test models and methods

### 3.1 Test methods

We chose two conventional black box testing methods to test the Wysa app, that is decision table, category partition testing methods. The decision table focuses more on the overall quality of the conversation while category partition testing methods focus more on testing the input data.

The decision table is a technique that tests the software behavior based on different combinations of input and the current state of the software. It is also called cause and effect table testing.

Category partition testing methods is a technique that divides input domains into different classes and each test case is taken from this class.

Scenario testing is about testing the app through the user's perspective and going through the whole scenario instead of just testing inputs.

### **3.2 Decision table testing**

Decision table testing is used to test the overall quality of the conversation, to see whether or not the chatbot can fulfill its three main functions: detect feelings, maintain conversations, and offer solutions to mental health problems.

### **3.3 Equivalence partitioning**

This type of test will be used to evaluate how good it is for Wysa to extract relevant info from diverse types of inputs, ranging from normal responses to invalid data inputs such as containing numbers only.

### **3.4 Scenario testing**

Scenario testing is a type of test to test how users would use the software end to end in different scenarios. We will use scenario testing to test different scenarios of how the user might use the app from start to finish. The test will not test invalid inputs/outputs but rather about use cases and conversation flow.

## Section #4 – Test case design with test data

### 4.1 Decision table testing

The four conditions below are selected to be part of the decision table testing

1. Is your answer related to the current conversation?

Wysa should be able to carry out conversations with users and change topics accordingly.

2. Does the answer express feelings?

Wysa should be able to detect the user's feelings and respond appropriately.

3. Does the answer change what you said previously?

Wysa should be able to update the information it has about you.

4. Is the answer one of the offered responses?

In certain situations, Wysa offers users some of the offered responses. However, it should also understand answers that do not follow the prompt to accommodate flexibility.

Conditions 1 and 4 must be either False or both True because if the user's answer using one of the offered responses (4), then it means the user's answer is related to the current conversation.

Outcomes

1. Wysa should ask to clarify what has just been said
2. Wysa should offer solutions to de-escalate the user's feelings if the feelings are negative
3. Wysa should be able to update old data with the new one
4. Wysa should be able to continue the conversation in the appropriate direction with message personalization

#### 4.1.1 Detailed Decision Table

| Conditions | R1 | R2<br>(X) | R3 | R4<br>(X) | R5 | R6<br>(X) | R7 | R8<br>(X) | R9<br>(X) | R10 | R11<br>(X) | R12 | R13<br>(X) | R14 | R15<br>(X) | R16 |
|------------|----|-----------|----|-----------|----|-----------|----|-----------|-----------|-----|------------|-----|------------|-----|------------|-----|
| C1         | F  | F         | F  | F         | F  | F         | F  | T         | T         | T   | T          | T   | T          | T   | T          | T   |
| C2         | F  | F         | F  | F         | T  | T         | T  | F         | F         | F   | F          | T   | T          | T   | T          | T   |
| C3         | F  | F         | T  | T         | F  | F         | T  | T         | F         | F   | T          | T   | F          | F   | T          | T   |
| C4         | F  | T         | F  | T         | F  | T         | F  | T         | F         | T   | F          | T   | F          | T   | F          | T   |
| Actions    |    |           |    |           |    |           |    |           |           |     |            |     |            |     |            |     |
| O1         | X  |           |    |           |    |           |    |           |           |     |            |     |            |     |            |     |
| O2         |    |           |    |           | X  |           | X  |           |           |     |            |     |            |     |            | X   |
| O3         |    |           | X  |           |    |           | X  |           |           |     |            | X   |            |     |            | X   |
| O4         | X  | X         | X  | X         | X  | X         | X  | X         | X         | X   | X          | X   | X          | X   | X          | X   |

R1: Invalid answer that does not contribute anything significant to the conversation

R3: The user wants to update previously said information

R5: The user suddenly wants to express how they are feeling

R7: The user suddenly wants to express how they are feeling, which is different from what previously said

R10: The user is continuing the conversation with the chatbot, performing an exercise, engaging in conversations, etc, and answered with one of the offered responses

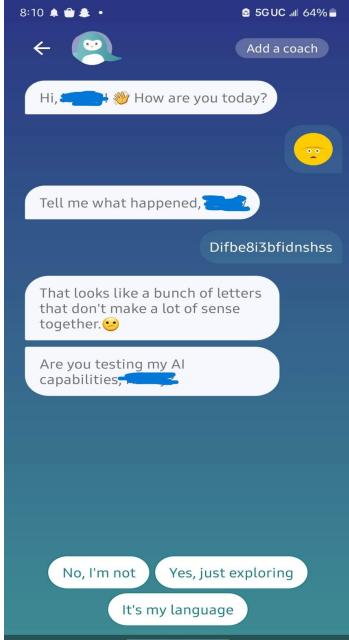
R12: The user is continuing the conversation with the chatbot, but wants to correct what was previously said and answered with one of the offered responses

R14: The user expressing how they feel to the chatbot as part of the current conversation and answered with one of the offered responses

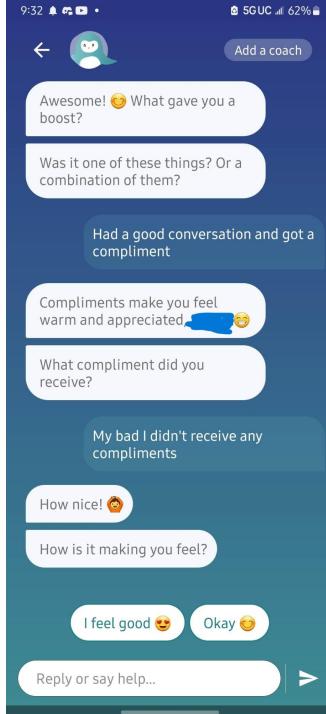
R16: The user expresses how they feel to the chatbot, but wants to correct what was previously said, maybe because their mood changes. User answered with one of the offered responses and answered with one of the offered responses

#### 4.1.2 Decision Table Test Cases

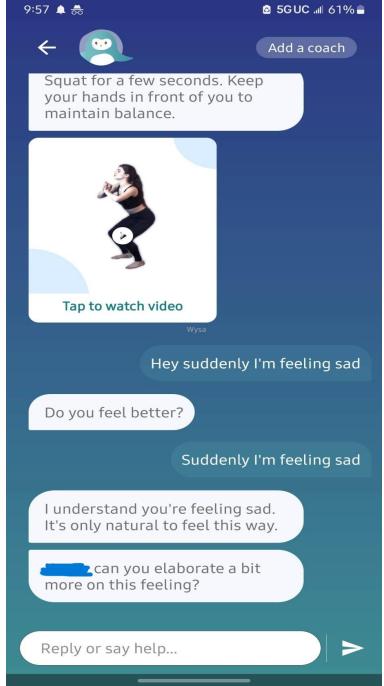
|                  |   |
|------------------|---|
| Test case ID     | R1  |
| Test Description | User inputs an invalid answer that does not contribute anything significant to the conversation |

|                  |   |
|------------------|---|
| Test case input  |  |
| Situation        | Incoherent text in the middle of the chat   |
| Expected Result  | Wysa should ask to clarify what just been said                                    |
| Actual Result    | Wysa said what entered didn't make any sense                                      |
| Test Case Result | Pass  |

|                  |  |
|------------------|--|
| Test case ID     | R3   |
| Test Description | User wants to update previously said information |

|                  |   |
|------------------|---|
| Test case input  |  <p>The screenshot shows a mobile application interface for Wysa. At the top, there's a header with a back arrow, a profile icon of a owl-like character, and a button labeled "Add a coach". Below the header, the main conversation area has the following messages:</p> <ul style="list-style-type: none"> <li>User: "Awesome! 😊 What gave you a boost?"</li> <li>Wysa: "Was it one of these things? Or a combination of them?"</li> <li>User: "Had a good conversation and got a compliment"</li> <li>Wysa: "Compliments make you feel warm and appreciated 😊"</li> <li>User: "What compliment did you receive?"</li> <li>Wysa: "My bad I didn't receive any compliments"</li> <li>User: "How nice! 🎉"</li> <li>Wysa: "How is it making you feel?"</li> <li>User: "I feel good 😊" (with a blue highlighted bubble)</li> <li>Wysa: "Okay 😊" (with a blue highlighted bubble)</li> </ul> <p>At the bottom of the screen, there's a text input field with the placeholder "Reply or say help..." and a right-pointing arrow button.</p> |
| Situation        | User said they were having a good day, and Wysa asked what happened. The user said they had a good conversation and had a compliment but later said those didn't happen.  |
| Expected Result  | Wysa should be able to update data with the new one and offer message personalization.  |
| Actual Result    | Wysa assumed that the text was part of the conversation and completely ignored it   |
| Test Case Result | Fail  |

|                  |   |
|------------------|---|
| Test case ID     | R5  |
| Test Description | User suddenly wants to express how they are feeling |

|                  |   |
|------------------|---|
| Test case input  |    |
| Situation        | Wysa was guiding the user through squat exercises then the user suddenly said they were feeling sad.  |
| Expected Result  | Wysa should offer de-escalation if there's any, as well as continue the conversation in the correct direction   |
| Actual Result    | Wysa was able to detect feelings inside the reply despite the whole conversation being about working out. Then it asks the user to elaborate on the feeling |
| Test Case Result | Pass  |

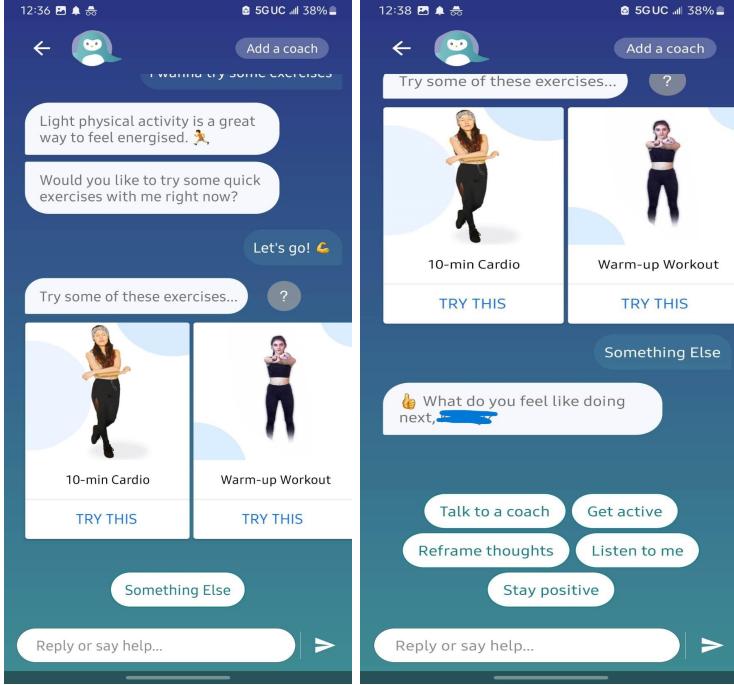
|                  |   |
|------------------|---|
| Test case ID     | R7  |
| Test Description | User suddenly wants to express how they are feeling, which is different from what previously said |

|                  |   |
|------------------|---|
| Test case input  |  <p>The screenshot shows a mobile application interface for Wysa. At the top, there's a small owl icon and a button to 'Add a coach'. Below that, there are two cartoon tooth characters. A message bubble from the user says 'What would you like to do next?'. Wysa's response is 'Hey I'm feeling happy'. Wysa then asks 'I'd love to hear more about what makes you feel happy.' Another message from the user asks 'Where all do you think this is stemming from?' Wysa replies 'Sorry I suddenly feeling sad'. Wysa then offers support with the message 'I understand. Becoming our own best friend can make us find that inner peace and alignment.' Finally, Wysa suggests 'Let's work on increasing our compassion towards ourselves.' The user responds with 'That sounds good'. At the bottom of the screen, there are two buttons: 'Start' and 'Maybe later'.</p> |
| Situation        | User suddenly expressed that they were happy, Wysa asked to elaborate on that feeling. However, the user suddenly changed their feeling and said they were feeling sad instead.   |
| Expected Result  | Wysa should offer de-escalation if there's any, update old data with old ones as well as continue the conversation in the correct direction and offer message personalization   |
| Actual Result    | Wysa failed to recognize the change in feelings and assume what the user said.  |
| Test Case Result | Fail  |

|                  |   |
|------------------|---|
| Test case ID     | R10   |
| Test Description | The user is continuing the conversation with the chatbot, performing an exercise, engaging in conversations, etc., and answered with one of the offered responses |

|                  |  |
|------------------|--|
| Test case input  |  |
| Situation        | Wysa asked the user how they were feeling. The user then proceeded with the conversations using one of the offered response choices and conversed with Wysa. |
| Expected Result  | Wysa should be able to continue the conversation in the correct direction with message personalization   |
| Actual Result    | Wysa was able to offer message personalization as well as continue the conversation  |
| Test Case Result | Pass   |

|                  |  |
|------------------|--|
| Test case ID     | R12  |
| Test Description | User is continuing the conversation with the chatbot, but wants to correct what was previously said using one of the offered responses |

|                         |   |
|-------------------------|---|
| <b>Test case input</b>  |  <p>The screenshots show a conversation between a user and a Wysa coach. In the first screenshot, the user asks for quick exercises, and the coach suggests a 10-min Cardio session. In the second screenshot, the user changes their mind, and the coach offers alternative options like 'Something Else' or 'Get active'.</p> |
| <b>Situation</b>        | User asked to go through a workout. After seeing a list of exercises, the user changed their mind   |
| <b>Expected Result</b>  | Wysa should be able to pick up the change of mind as well as continue the conversation in the correct direction   |
| <b>Actual Result</b>    | Wysa was able to notice the user didn't want to work out anymore and suggested other things to talk   |
| <b>Test Case Result</b> | Pass  |

|                         |  |
|-------------------------|--|
| <b>Test case ID</b>     | R14  |
| <b>Test Description</b> | User expresses how they feel to the chatbot as part of the current conversation using one of the offered responses |

|                        |   |
|------------------------|---|
| <b>Test case input</b> | <p>The screenshot shows a mobile application interface for a chatbot named 'Wysa'. At the top, there's a header with the time '1:03', signal strength, battery level '36%', and a '5G UC' indicator. Below the header is a profile picture of a cartoon owl and a button 'Add a coach'. The main area is a conversation screen. The user's message 'Hi, Ok! 😊 How are you today?' is at the top. The bot's response 'Most days we are somewhere in the middle of our emotional range.' follows. A question from the user 'If you were to try and describe your current mood, what kind of an 'okay' are you feeling right now?' is shown. Below these messages is a row of six circular emotion buttons: 'Meh', 'Optimistic', 'Calm', 'Numb', 'Bored', and 'Confused'. At the bottom is a text input field with 'Reply or say help...' and a send button with a right-pointing arrow.</p> |
| Situation              | Wysa asked for the user's feelings whenever a new chat started. The user responded with an adjustable face feeling ok.  |
| Expected Result        | Wysa should be able to track user's feelings and offer an appropriate response  |
| Actual Result          | Wysa detected the user feeling ok and asked the user if they were ready to elaborate  |
| Test Case Result       | Pass  |

|                     |  |
|---------------------|--|
| <b>Test case ID</b> | R16  |
| Test Description    | User expresses how they feel to the chatbot, but wants to correct what was previously said, maybe because their mood change using one of the offered responses |

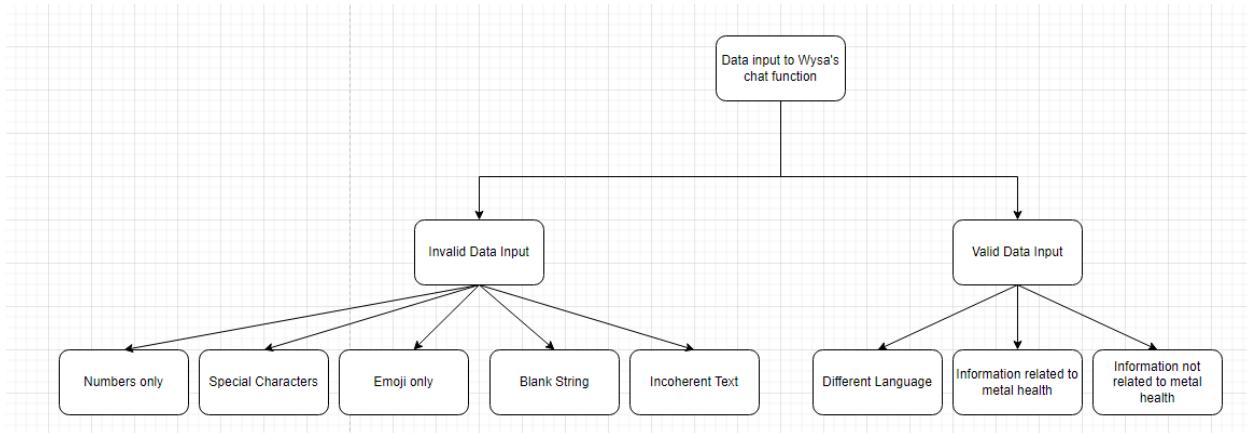
|                         |  |  |
|-------------------------|--|--|
| <b>Test case input</b>  | <p>Move the smiley up or down</p>  | <p>There's more Not now</p> <p>Reply or say help... &gt;</p> |
| <b>Situation</b>        | Wysa asked for the user's feelings when a new chat started. User responded with a sad face. After elaborating, Wysa offered ways to improve the situation. The user responded with one of the offered responses and said they felt better. |  |
| <b>Expected Result</b>  | Wysa was able to notice the change in feelings and continue the conversation in the correct conversation.  |  |
| <b>Actual Result</b>    | Wysa was able to notice that the user was feeling better and wished the user a good day  |  |
| <b>Test Case Result</b> | Pass   |  |

### 4.1.3 Decision Table Test Coverage

Even though we have 16 combinations, 8 of the combinations didn't make sense, as explained above: condition 1 and 4 must be both true or both false, because if the user answers using one of the responded questions, then the user is following the prompt. That left us with 8 combinations left, and since we have one test case for each partition, this means we have 100% coverage.

## 4.2 Equivalence testing

### 4.2.1 Detailed Equivalence Partitioning Tree



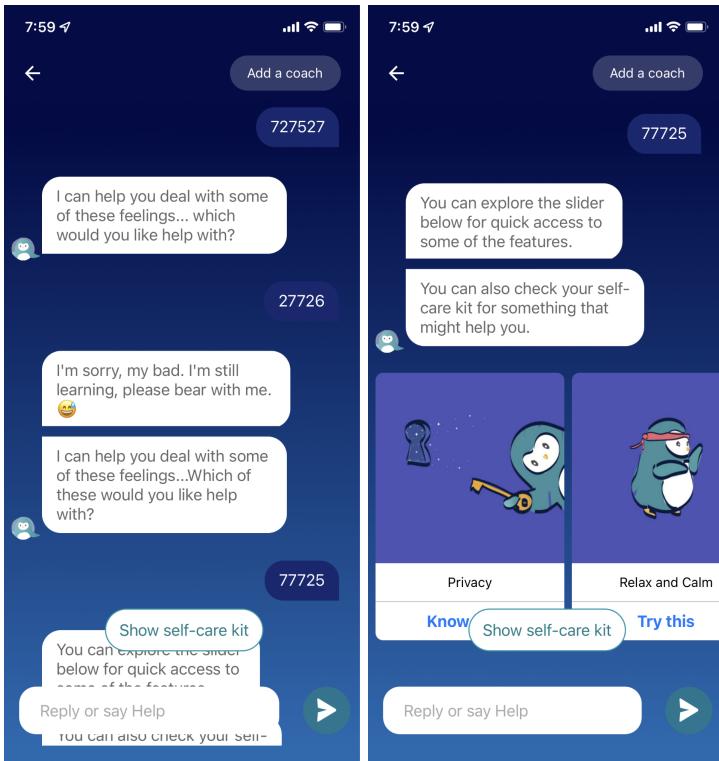
#### 1) Invalid Inputs

- a) Respond in numbers
- b) Respond in incoherent texts
- c) Respond in special characters
- d) Respond in emojis
- e) Respond in blank strings

#### 2) Valid Inputs

- a) Respond in different languages
- b) Provide information related to your mental health
- c) Provide information not related to your mental health

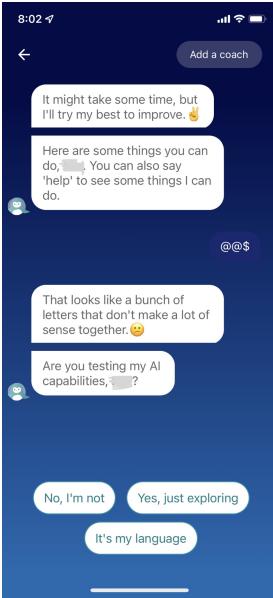
#### 4.2.2 Equivalence Partitioning Test Cases

|                  |   |
|------------------|---|
| Test Case ID     | 1   |
| Test Description | Respond in numbers only   |
| Test Input       |  |
| Expected Result  | Tell the user it does not understand what was said                                  |
| Actual Result    | Kept recommending the app's various services/features                               |
| Test Case        | Failed  |

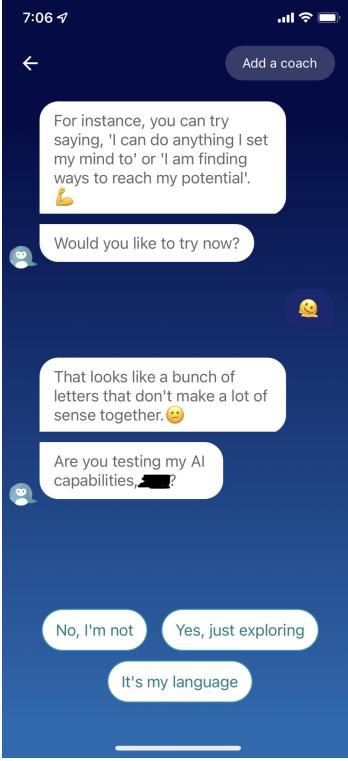
|        |  |
|--------|--|
| Result |  |
|--------|--|

|                  |  |
|------------------|--|
| Test Case ID     | 2  |
| Test Description | Respond in incoherent texts  |
| Test Input       |  |
| Expected Result  | Carry on as normal   |
| Actual Result    | Asks if you are testing the AI's capabilities                                      |
| Test Case Result | Passed   |

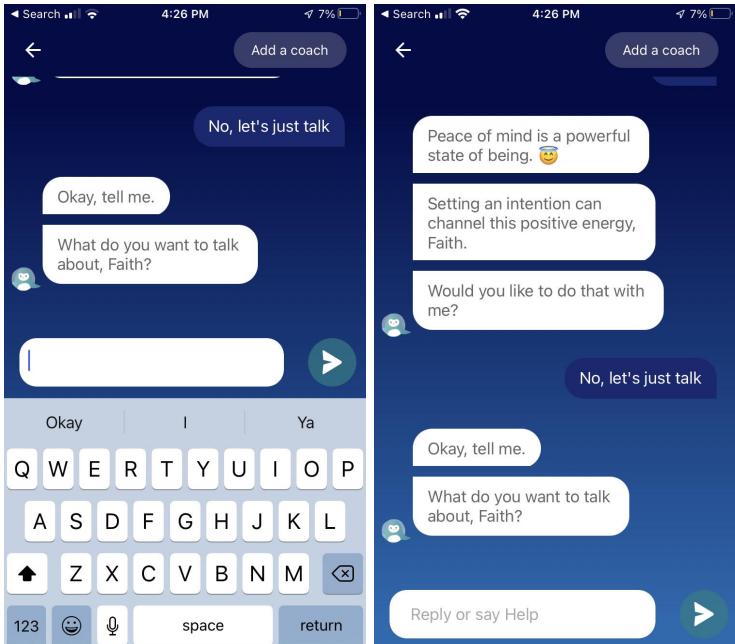
|                  |                               |
|------------------|-------------------------------|
| Test Case ID     | 3                             |
| Test Description | Respond in special characters |

|                  |  |
|------------------|--|
| Test Input       |  <p>The screenshot shows a mobile application interface for a coaching service. At the top, it says "Add a coach". The conversation starts with the AI coach saying:</p> <p>"It might take some time, but I'll try my best to improve. 🌟"</p> <p>Then, the user types "@@\$". The AI coach responds:</p> <p>"Here are some things you can do, @@\$. You can also say 'help' to see some things I can do."</p> <p>Next, the user types "That looks like a bunch of letters that don't make a lot of sense together. 😊". The AI coach replies:</p> <p>"Are you testing my AI capabilities, 😊?"</p> <p>Finally, the user types "No, I'm not" and "Yes, just exploring". The AI coach concludes with:</p> <p>"It's my language"</p> |
| Expected Result  | Ask what the user typed  |
| Actual Result    | Asks if you are testing the AI's capabilities  |
| Test Case Result | Passed   |

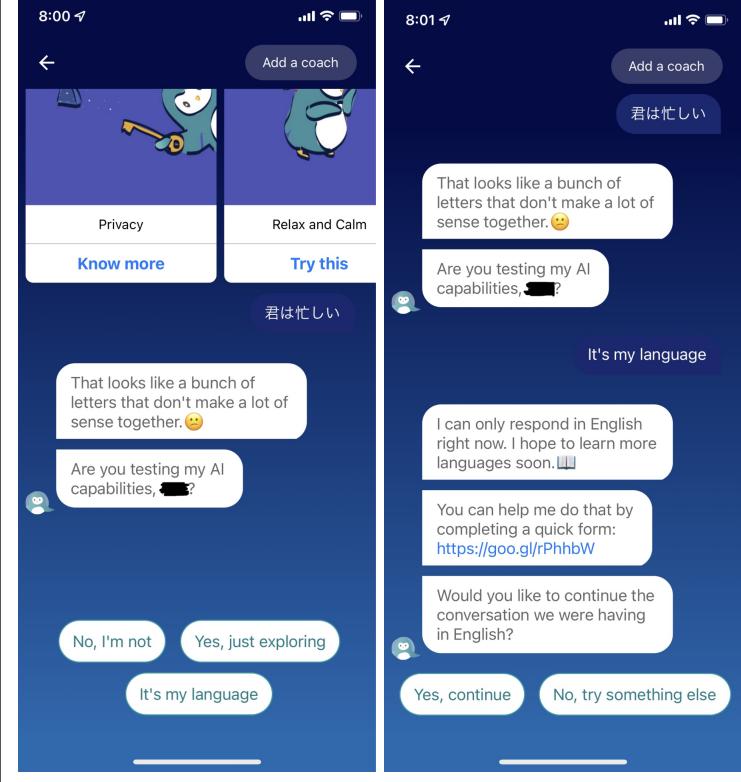
|                  |                   |
|------------------|-------------------|
| Test Case ID     | 4                 |
| Test Description | Respond in emojis |

|                  |  |
|------------------|--|
| Test Input       |  <p>The screenshot shows a mobile application interface. At the top, it says "7:06" and has signal, battery, and notification icons. There is a back arrow and a button labeled "Add a coach". The main area is a conversation between a user and a coach. The user's messages are in blue bubbles, and the coach's responses are in white bubbles with a dark blue background. The user's first message is: "For instance, you can try saying, 'I can do anything I set my mind to' or 'I am finding ways to reach my potential'. 🤸". The coach's response is: "Would you like to try now? 😊". The user's second message is: "That looks like a bunch of letters that don't make a lot of sense together. 😐". The coach's response is: "Are you testing my AI capabilities? 🤖". At the bottom, there are three options: "No, I'm not", "Yes, just exploring", and "It's my language".</p> |
| Expected Result  | Understands most of the face emojis  |
| Actual Result    | Asks if you are testing the AI's capabilities  |
| Test Case Result | Passed   |

|                  |                          |
|------------------|--------------------------|
| Test Case ID     | 5                        |
| Test Description | Respond in blank strings |

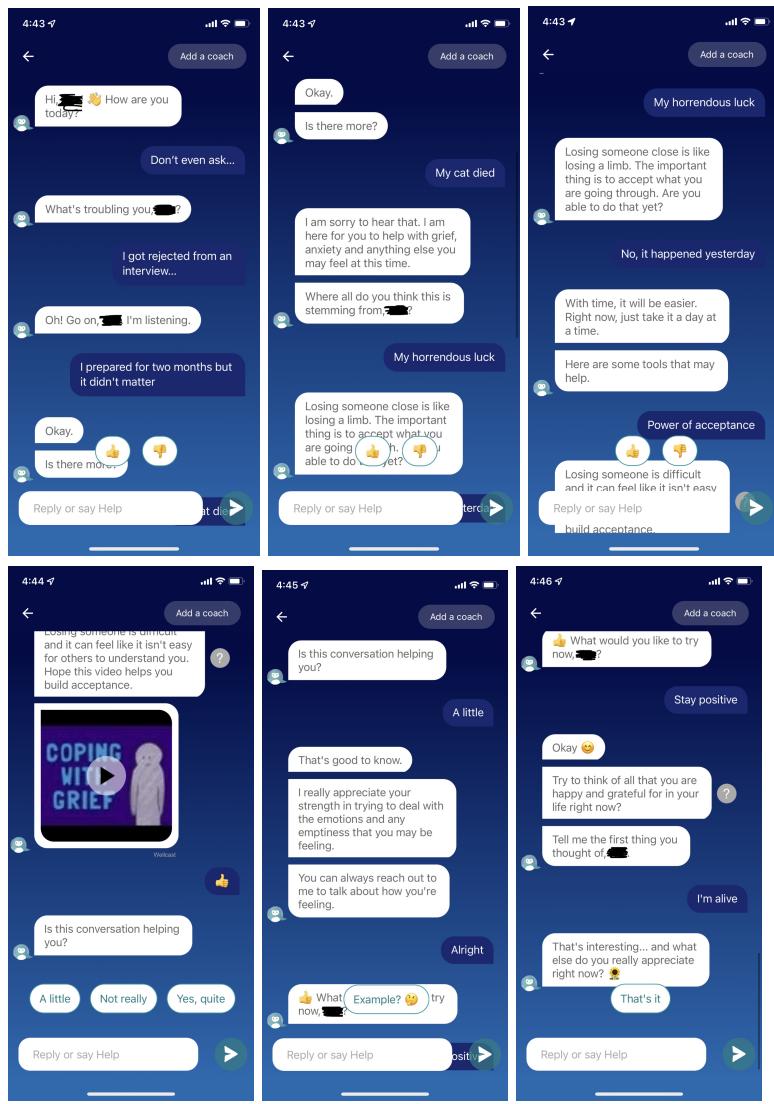
|                   |  |
|-------------------|--|
| <b>Test Input</b> |  |
| Expected Result   | Not allowed/ask the user to type something in the chat bar                         |
| Actual Result     | Not allowed, doesn't allow sending a blank space only                              |
| Test Case Result  | Passed   |

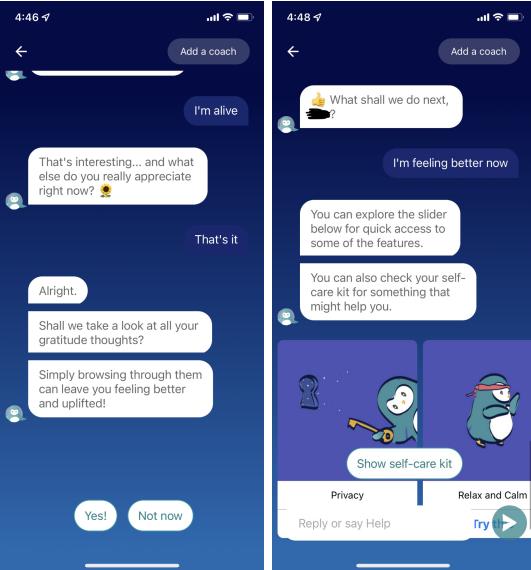
|                  |                                |
|------------------|--------------------------------|
| Test Case ID     | 6                              |
| Test Description | Respond in different languages |

|                  |  |
|------------------|--|
| Test Input       |  <p>The screenshots show a mobile application interface. The left screenshot is at 8:00 and the right is at 8:01. Both screens have a dark blue background. At the top, there's a navigation bar with a back arrow and a button labeled "Add a coach". Below this, there are two cartoon owl icons. The left owl is associated with a "Privacy" card and a "Know more" button. The right owl is associated with a "Relax and Calm" card and a "Try this" button. In the center, there's a message input field with placeholder text "君は忙しい" (You are busy). Below the input field, there are two speech bubbles from the user:</p> <ul style="list-style-type: none"> <li>"That looks like a bunch of letters that don't make a lot of sense together. 😊"</li> <li>"Are you testing my AI capabilities, [REDACTED]?"</li> </ul> <p>Below these messages, there are two buttons: "No, I'm not" and "Yes, just exploring". Underneath these buttons is another message bubble from the user: "It's my language".</p> <p>In the right screenshot (8:01), the AI response is shown in a speech bubble:</p> <ul style="list-style-type: none"> <li>"That looks like a bunch of letters that don't make a lot of sense together. 😊"</li> <li>"Are you testing my AI capabilities, [REDACTED]?"</li> <li>"It's my language"</li> <li>"I can only respond in English right now. I hope to learn more languages soon. 📚"</li> <li>"You can help me do that by completing a quick form: <a href="https://goo.gl/rPhhbW">https://goo.gl/rPhhbW</a>"</li> <li>"Would you like to continue the conversation we were having in English?"</li> <li>"Yes, continue" and "No, try something else"</li> </ul> |
| Expected Result  | Would you like to switch languages? Immediate response in used language  |
| Actual Result    | Does not understand languages other than English   |
| Test Case Result | Passed   |

|                  |   |
|------------------|---|
| Test Case ID     | 7   |
| Test Description | Provide information related to your mental health |

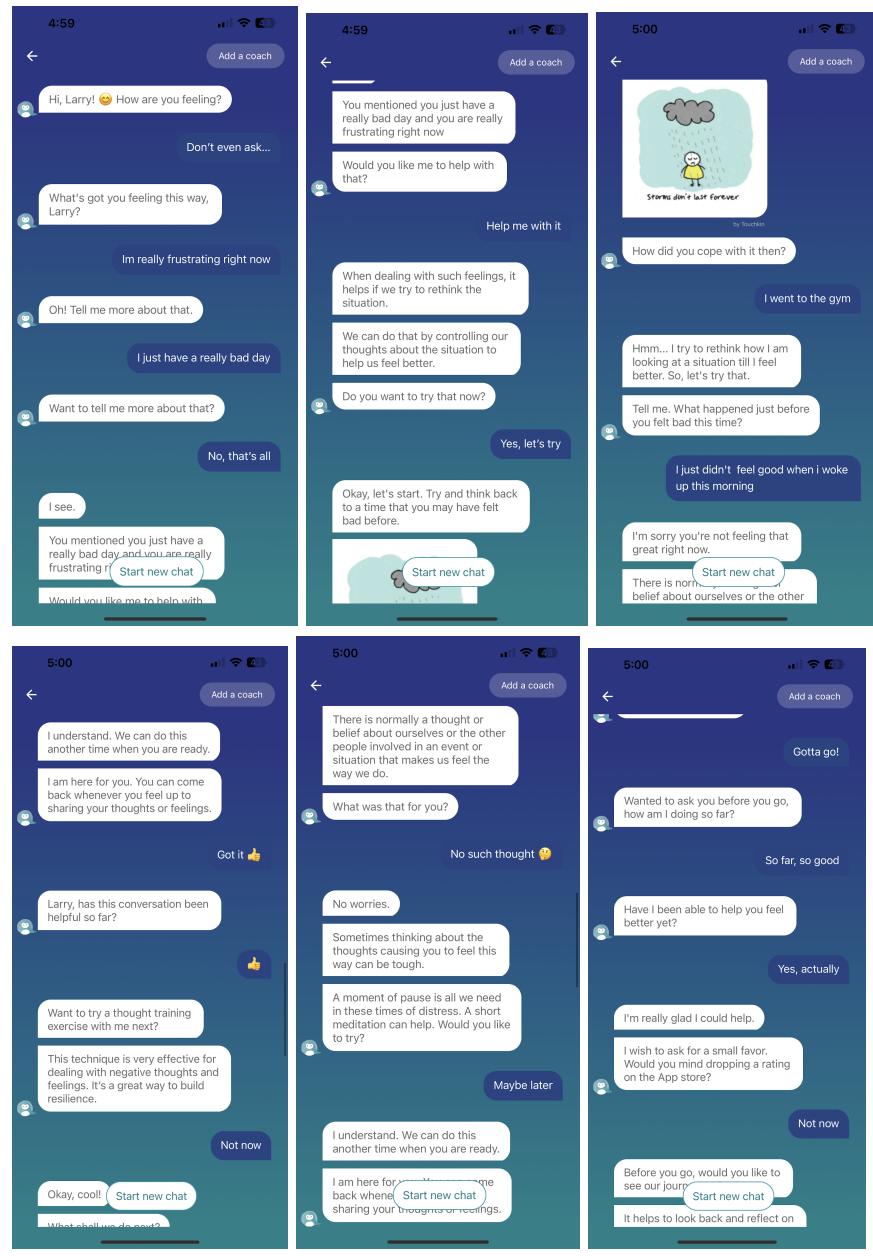
## Test Input

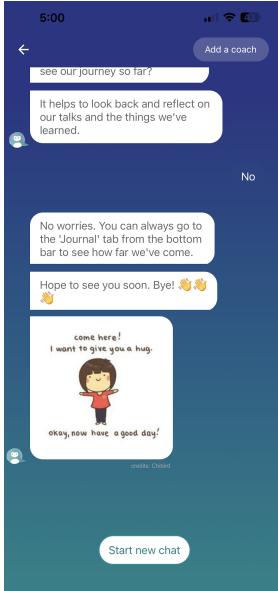


|                  |  |
|------------------|--|
|                  |                                 |
| Expected Result  | The app will try to help if you are in a negative mindset, or if you are feeling positive get you to explain why |
| Actual Result    | Attempted to talk through what was bothering the user and provided potential exercises to help cope with loss    |
| Test Case Result | Passed   |

|                  |   |
|------------------|---|
| Test Case ID     | 8   |
| Test Description | Provide information not related to your mental health |

## Test Input



|                  |  |
|------------------|--|
|                  |                               |
| Expected Result  | The app will ask the user to keep talking and try to relate it to how you feel                                 |
| Actual Result    | The app will ask the user to keep talking and try to help the users with some thought processes and exercises. |
| Test Case Result | Passed   |

#### 4.2.3 Equivalence Partitioning Test Coverage

There are 8 test cases for the equivalence partitioning test. This test will guarantee each partition of the data input is covered with 1 test case supporting that partition.

For partitions of valid input data for the chat function, these partitions will include messages with topics related to mental health, topics not related to mental health, or in different languages. This test will aim to ensure that the chat function could handle all valid inputs correctly and give the expected results.

For partitions of invalid input data, these partitions will be messages with a blank string, emojis, incoherent text, special characters, or numbers. These partitions will ensure that the chat function of Wysa will handle error messages and invalid inputs.

## 4.3 Scenario testing

### 4.3.1 Scenario Testing Designs

There are four scenarios that will be tested:

1. User wants to use the app to look for SOS functions
2. User wants to use the app to express they just have a bad day
3. User wants to use the app to express they have a good day
4. User wants to use the app to practice mindfulness activities

The thing with chatbots is that it is hard to stop an ongoing conversation with it, especially Wysa. Therefore the scenario can stop when the conversation is finished or done

#### 4.3.1.1 SOS Scenario

| Scenario 1 | Panic Crisis Aid  |
|------------|---|
| 1          | User opens app  |
| 2          | User presses on “SOS”   |
| 3          | User presses on “Grounding for panic”                                   |
| 4          | Wysa asks the user if they can suggest some exercise to help calm down? |
| 6          | User chooses “Yes”  |
| 7          | Wysa suggests exercises for the user to follow                          |
| 8          | User does suggested exercises   |

|            |   |
|------------|---|
| 9          | User ends the conversation when done                                  |
| Scenario 2 | Safety Plan   |
| 1          | Wysa asks the User to list events that causes anxiety or stress       |
| 2          | User lists down events  |
| 3          | Wysa asks the User to list everyone that can offer safety in a crisis |
| 4          | User list down nicknames  |
| 5          | Wysa asks User to list safe places                                    |
| 6          | User lists down places/locations                                      |
| 7          | Wysa asks User to list activities that provides calm or comfort       |
| 8          | User lists down activities  |
| 9          | Wysa creates a safety plan  |
| 10         | User checks safety plan   |
| Scenario 3 | Crisis Services   |
| 1          | User opens app  |
| 2          | User presses on “SOS”   |
| 3          | User presses on “International Crisis Helplines”                      |
| 4          | User chooses their country location                                   |
| 5          | Wysa lists a list of crisis helplines                                 |
| 6          | User presses the back button  |
| 3          | User presses on “International Child Crisis Helplines”                |
| 4          | User chooses their country location                                   |
| 5          | Wysa lists a list of crisis helplines                                 |

#### 4.3.1.2 Bad Day Scenario

|            |                          |
|------------|--------------------------|
| Scenario 1 | User is having a bad day |
|------------|--------------------------|

|   |  |
|---|--|
| 1 | User opens app                                   |
| 2 | User starts chatting with Wysa                   |
| 3 | User expresses that they just had a bad day      |
| 4 | User went through conversation with Wysa         |
| 5 | The scenario ended when the conversation is done |

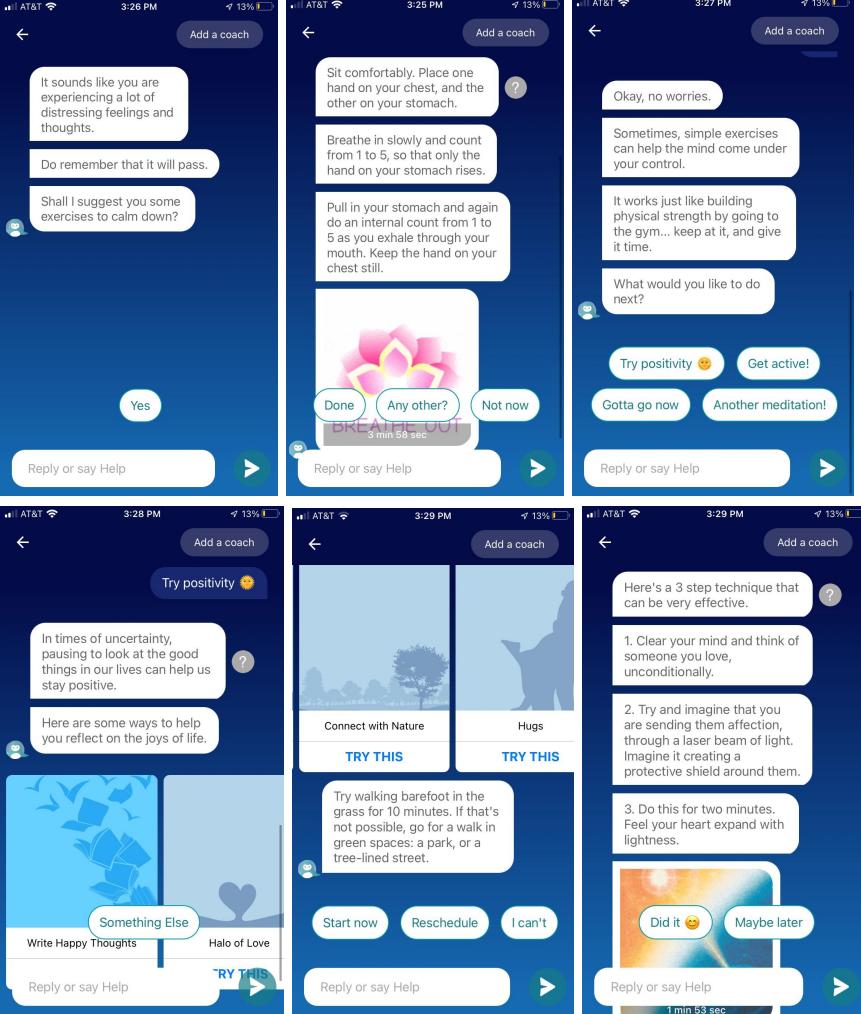
#### 4.3.1.3 Good day Scenario

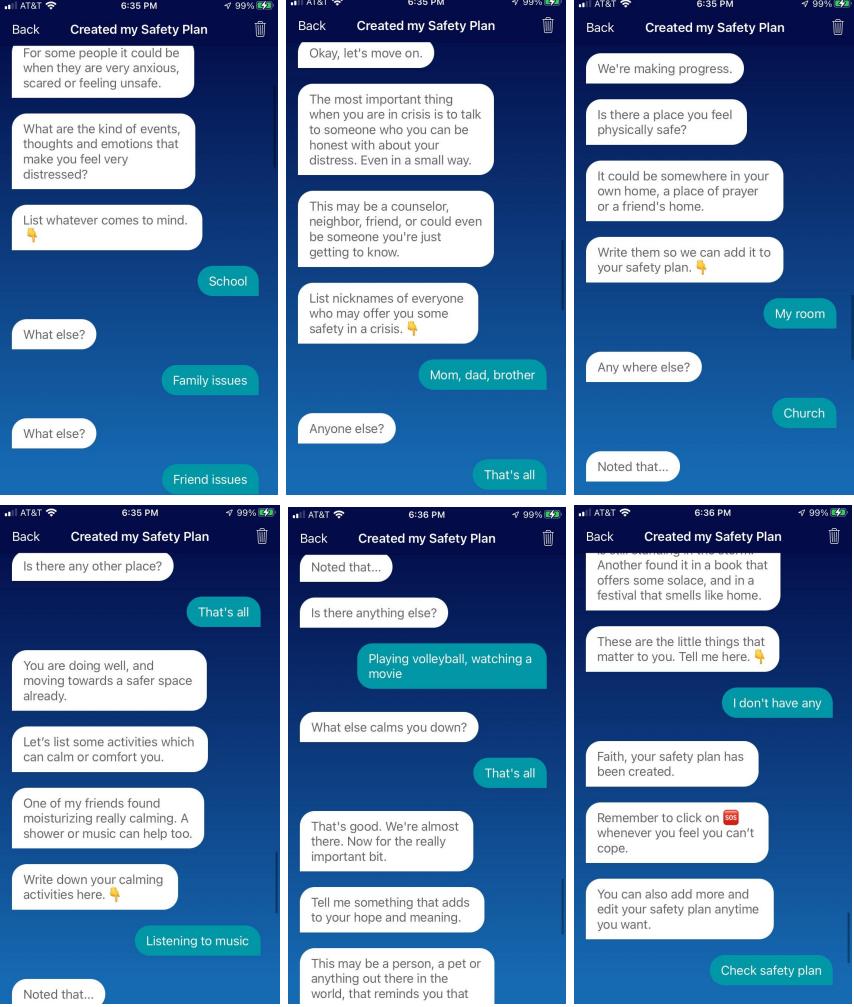
|            |  |
|------------|--|
| Scenario 1 | User is having a good day                        |
| 1          | User opens app                                   |
| 2          | User starts chatting with Wysa                   |
| 3          | User expresses that they just had a good day     |
| 4          | User went through conversation with Wysa         |
| 5          | The scenario ended when the conversation is done |

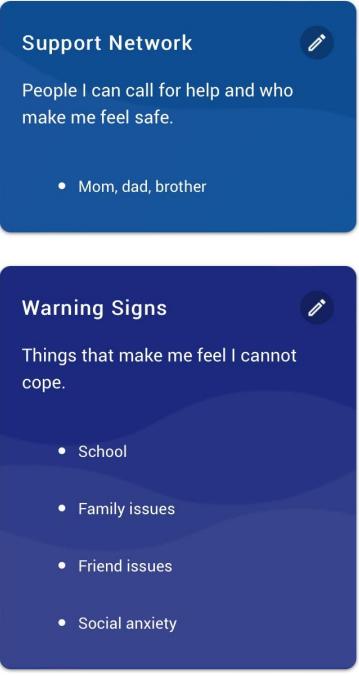
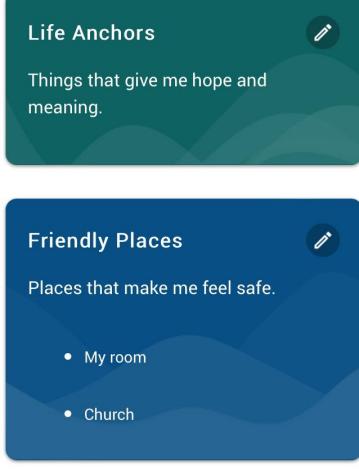
#### 4.3.1.4 Mindfulness Activities Scenario

|            |  |
|------------|--|
| Scenario 4 | User wants to use the app to practice mindfulness activities |
| 1          | User opens app   |
| 2          | User starts chatting with Wysa                               |
| 3          | User expresses wanting to practice mindfulness activities    |
| 4          | User goes through mindfulness activities with Wysa           |
| 5          | User ends the conversation when done                         |

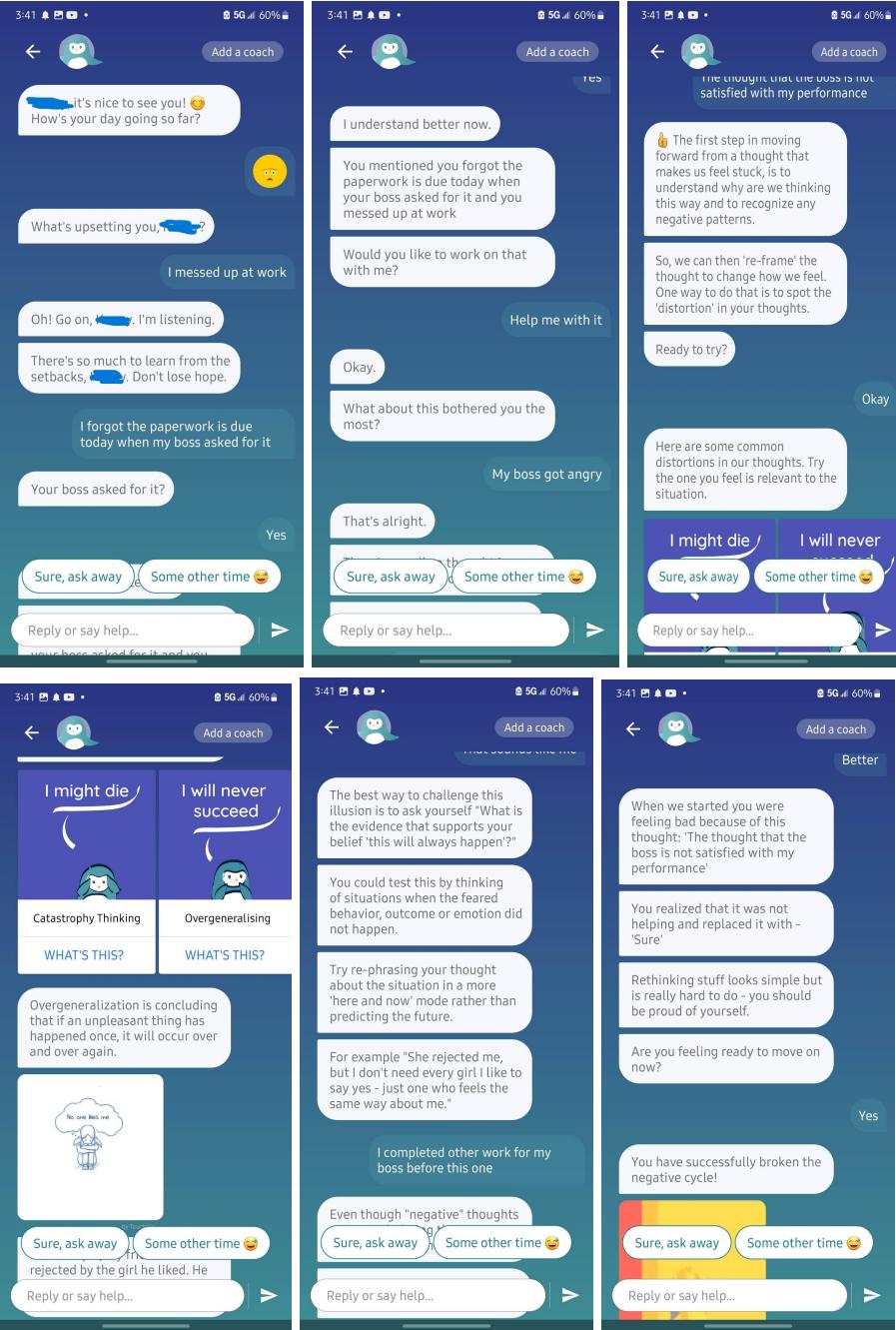
### 4.3.2 Scenario Testing Test Cases

|                  |  |
|------------------|--|
| Test Preparation | User is in SOS Options   |
| Test case ID     | 1.1  |
| Test case name   | User uses the “Grounding for panic” function   |
| State            | SOS  |
| Test details     |  <p>The screenshots illustrate a user interaction in the SOS mode of the Wysa app. The user starts by asking for grounding exercises, which are then provided. The user then chooses the 'Try positivity' option, leading to a series of three recommended activities: 'Connect with Nature', 'Hugs', and 'TRY THIS' (walking barefoot). Each activity card includes a 'TRY THIS' button and a timer.</p> |
| Situation        | User is experiencing distressing feelings or thoughts.   |
| Expected Result  | Wysa should be able to offer mental aid that will help calm the User until the crisis resolves.  |
| Actual Result    | Wysa was able to offer relaxing exercises and positive activities.   |

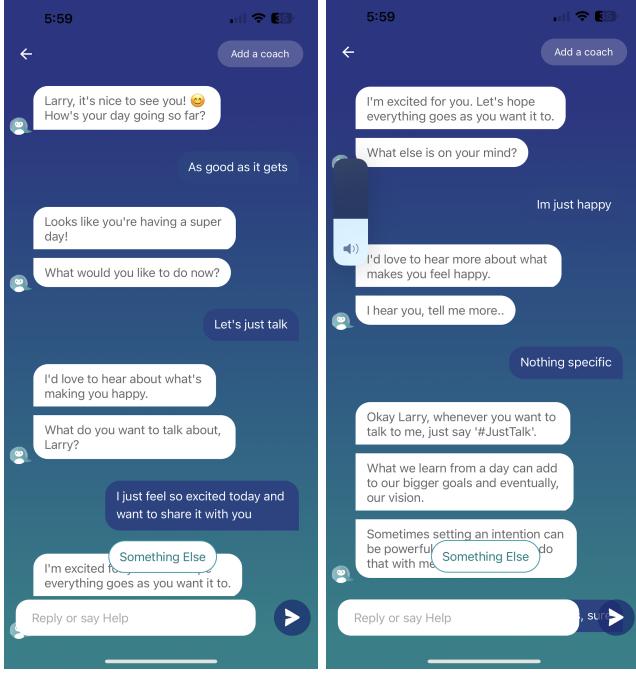
|                 |   |
|-----------------|---|
| Test Case ID    | 1.2   |
| Test Case Name  | User Creates a Safety Plan  |
| State           | SOS   |
| Test Details    |  <p>The screenshots show a conversational interface between the user and the AI. The user asks questions and provides responses, which the AI uses to generate a safety plan. The AI suggests various coping mechanisms and safe spaces based on the user's input.</p> |
| Situation       | The User would like to create a safety plan in case of a future mental or emotional crisis.   |
| Expected Result | Wysa should be able to create a safety plan that contains resources and support for the user in case of a crisis.   |
| Actual Result   | Wysa was able to create a safety plan for the User that contains a list of professional support and helplines. Based on the user's response, Wysa was able to create a list of life anchors, friendly places, support networks, warning signs, and calming activities for the User.   |

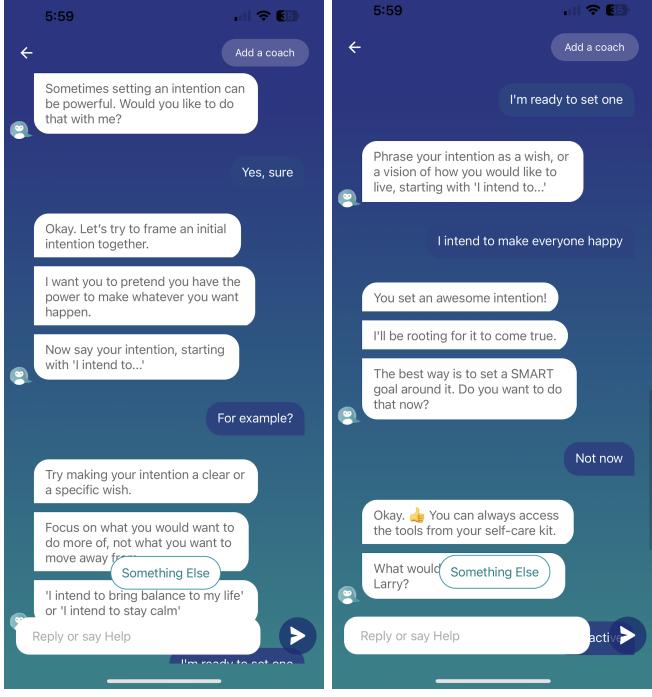
|        |   |  |
|--------|---|--|
|        |    |    |
|        |  |  |
| Result | Test case passed  |  |

|              |   |
|--------------|---|
| Test case ID | 2 |
|--------------|---|

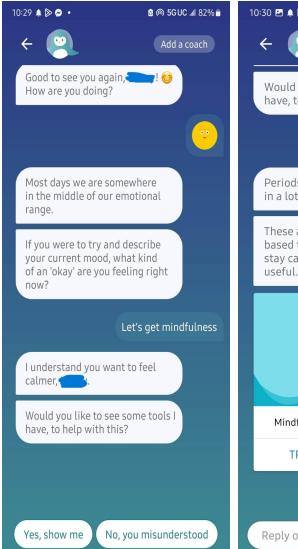
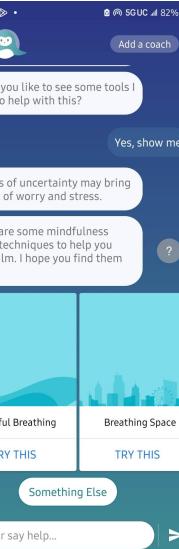
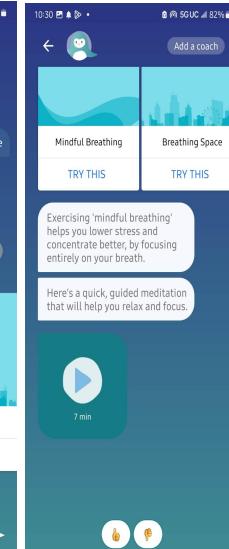
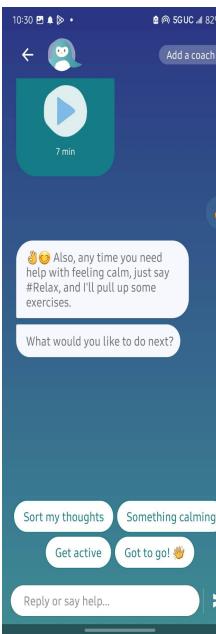
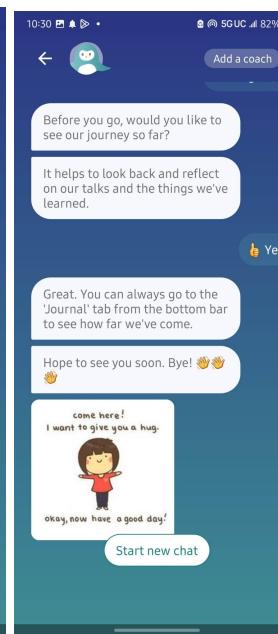
|                  |  |
|------------------|--|
| Test Preparation | User is in a chat conversation   |
| Test case name   | User wants to use the app to express they have a bad day   |
| State            | Detected bad feeling   |
| Test details     |  <p>The screenshots show a user interaction with the Wysa app. The user expresses they have a bad day, mentioning being upset at work due to a boss's request. The AI coach responds with empathy and asks questions to understand better. The user identifies their boss as angry. The AI coach then guides the user through identifying negative thoughts, specifically catastrophizing and overgeneralizing. It provides examples and rephrasing techniques to challenge these thoughts, ultimately helping the user break the cycle of negative thinking.</p> |
| Situation        | User expressed that they had a bad day, Wysa guided the user through some mind activities in order to break the negative cycle   |

|                 |   |
|-----------------|---|
| Expected result | Wysa should be able to offer some solutions to make the user feels better and improve the situation |
| Actual result   | Wysa offered some thoughts solution to make the user feels better                                   |
| Result          | Test case passed  |

|                  |  |
|------------------|--|
| Test case ID     | 3  |
| Test Preparation | User is in a chat conversation   |
| Test case name   | Users want to express how good they feel   |
| State            | Detected good feeling  |
| Test details     |  <p>The image contains two side-by-side screenshots of a mobile application interface, likely the Wysa app. Both screenshots show a dark-themed chat screen with a blue header bar. The top bar includes a back arrow, a 'Add a coach' button, and signal strength indicators. The left screenshot shows a message from a user: 'Larry, it's nice to see you! 😊 How's your day going so far?'. Below it, a Wysa response says: 'As good as it gets'. Another user message follows: 'Looks like you're having a super day!'. Wysa responds: 'What would you like to do now?'. The user replies: 'Let's just talk'. Wysa then asks: 'I'd love to hear about what's making you happy.' The user responds: 'What do you want to talk about, Larry?'. Wysa replies: 'I just feel so excited today and want to share it with you'. The user then says: 'Something Else'. Wysa responds: 'I'm excited to hear that everything goes as you want it to.' At the bottom of the left screen, there is a text input field with 'Reply or say Help' and a blue send button with a white arrow. The right screenshot continues the conversation. The user says: 'I'm excited for you. Let's hope everything goes as you want it to.' Wysa responds: 'What else is on your mind?'. The user replies: 'Im just happy'. Wysa asks: 'I'd love to hear more about what makes you feel happy.' The user says: 'I hear you, tell me more..'. Wysa replies: 'Nothing specific'. The user then says: 'Okay Larry, whenever you want to talk to me, just say '#JustTalk'. Wysa responds: 'What we learn from a day can add to our bigger goals and eventually, our vision.' The user says: 'Sometimes setting an intention can be powerful'. Wysa replies: 'Something Else'. At the bottom of the right screen, there is a text input field with 'Reply or say Help' and a blue send button with a white arrow.</p> |

|                 |   |
|-----------------|---|
|                 |                       |
| Situation       | User have a good day and want to share it with Wysa   |
| Expected Result | Wysa should be able to keep the conversation going  |
| Actual Result   | Wysa kept the conversation going by providing some exercises and kits to make the user feel even better |
| Result          | Test Case Passed  |

|                  |  |
|------------------|--|
| Test case ID     | 4  |
| Test Preparation | User is in a chat conversation                               |
| Test case name   | User wants to use the app to practice mindfulness activities |
| State            | Guiding through mindfulness activities                       |

| Test details    |  |
|-----------------|--|
|                 |    |
|                 |     |
| Situation       | User asked Wysa for mindfulness activities, and Wysa recommended some for the user to choose from. Then Wysa went through the exercise in detail. After that, Wysa wished the user a good day  |
| Expected result | Wysa should be able to offer some message personalization regarding the situation  |
| Actual result   | Wysa offered a variety of mindfulness activity and guided user through the exercise  |
| Result          | Test case passed   |

### 4.3.3 Scenario Testing Coverage

For each scenario listed, we have at least one test case to cover. However, for the SOS function, we have 2 different test cases to test different aspects of the function. Therefore we can say we have 100% test coverage for scenario testing.

## Section #5 – Test result analysis and bug summary

### 5.1 Test Results Summary

#### Decision Table Test Summary

| Test case   | Result |
|---|--------|
| R1: Invalid answer that does not contribute anything significant to the conversation  | Passed |
| R3: The user wants to update previously said information  | Failed |
| R5: The user suddenly wants to express how they are feeling   | Passed |
| R7: The user suddenly wants to express how they are feeling, which is different from what previously said   | Failed |
| R10: The user is continuing the conversation with the chat bot, performing an exercise, engaging in conversations, etc and answered with one of the offered responses | Passed |

|   |        |
|---|--------|
| R12: The user is continuing the conversation with the chat bot, but wants to correct what was previously said and answered with one of the offered responses  | Passed |
| R14: The user expresses how they feel to the chatbot as part of the current conversation and answered with one of the offered responses   | Passed |
| R16: The user expresses how they feel to the chatbot, but wants to correct what was previously said, maybe because their mood changes. User answered with one of the offered responses and answered with one of the offered responses | Passed |

Pass percentage: 6/8 = 75%

### Equivalence Partitioning Test Summary

| Test Case                                | Result |
|--|--------|
| Numbers Only                             | Failed |
| Special Characters Only                  | Passed |
| Emojis                                   | Passed |
| Incoherent Text                          | Passed |
| Blank string                             | Passed |
| Inputs in different languages            | Passed |
| Information related to mental Health     | Passed |
| Information not related to mental Health | Passed |

The pass percentage is  $7/8 = 87.5\%$

### Scenario testing test summary

| Scenario                      | Result |
|-------------------------------|--------|
| SOS Scenario                  | Passed |
| Bad day scenario              | Passed |
| Good day scenario             | Passed |
| Mindfulness activity scenario | Passed |

Pass percentage is  $4/4 = 100\%$

## 5.2 Manual Test Costs

| Test                  | Cost        |
|-----------------------|-------------|
| Decision table        | 360 minutes |
| Equivalence partition | 300 minutes |
| Scenario testing      | 240 minutes |

## 5.3 Complexity and coverage

- Decision table: 5 test cases total
- Equivalence partitioning: 8 test cases total
- Scenario testing: 8 test cases total

## 5.4 Bug analysis and summary

### Bugs

- Wysa often assumes what user is saying when it asks user to elaborate on things or middle of the conversation

## Detecting feelings

- Wysa is good at detecting feelings in sentences as well when users change their feelings and can offer help.
- Better when the user is answered using one of the offered responses, if not then it is better for the user to keep their sentences as concise as possible

## Conversation flow

- Wysa does not offer much conversation variety
- Wysa often assumes what the user saying when the user is in the middle of the conversation
- To change the subject, it is better to end the current subject first or signal Wysa that you want to say something else instead of doing it abruptly

## Help offering

- Wysa has a good selection of exercises to offer help
- Can guide the user through exercise in details
- Ask for feedback whether or not the exercise was good enough

## Invalid Inputs

- For most invalid inputs, Wysa will ask the users if they are testing the AI's capability.
- Wysa will recommend its various features or services.
- Wysa will not allow blank strings.

## Valid Inputs

- Wysa will not understand any other language than English and will ask users to either continue with English or stop the conversation.

- Wysa will attempt to keep the conversation going even if the messages are not related to mental health and provide some potential exercises for how the users feel.

#### Conversation flow

- Wysa was able to support its main functions
  - For SOS functions, Wysa guides user through deescalation exercises as well as having support for creating Safety Plan
  - When Wysa detects that user is in a bad mood, it offers thought exercises to reframe the mind to stop the negative thinking
  - When Wysa detects that user is in a good mood, it offers the user to elaborate on it
  - If a user asks to go through some mindfulness activities, Wysa will then offer a variety of them and guide the user through the chosen one.

**Deliverable 2B - AI Test Report**

**CMPE 187 - Software Quality Engineering**

**San Jose State University**

**Group 2**

Team members:

Khai Nguyen

Faith Deanon

Kevin Campbell

Long Nguyen

|  |           |
|--|-----------|
| <b>Deliverable #2)</b>   | <b>6</b>  |
| Section 2.1 Context Modeling for AI Features                           | 7         |
| Mental Illness function  | 7         |
| Exercise function  | 7         |
| Feeling Detection function   | 8         |
| Invalid Topics Function  | 10        |
| Section 2.2 AI Function Input Classifications                          | 11        |
| Mental Illness function  | 11        |
| Exercise function  | 11        |
| Invalid Topics Function  | 13        |
| Feeling Detection Function   | 13        |
| Section 2.3 AI Function Output/Event/Action Classifications            | 15        |
| Mental Illness function  | 15        |
| Exercise function  | 15        |
| Invalid Topics Function  | 16        |
| Feeling detection function   | 16        |
| Section 2.4 AI Function Classification Decision Table (3D tables)      | 18        |
| Mental Illness function  | 18        |
| Exercise function  | 19        |
| Invalid Topics Function  | 21        |
| <b>Section 3 – AI Function Test Cases with Inputs/Expected Outputs</b> | <b>26</b> |
| Section 3.1 Test data models   | 26        |
| Feeling Detection function   | 26        |
| Mental Illness function  | 36        |
| Decision Table   | 36        |
| Test case 1: Recommends an exercise/technique                          | 37        |
| Test case 2: Change in User's conversation format                      | 38        |
| Test case 3: Short Term Memory   | 39        |
| Test case 4: Long Term Memory  | 40        |
| Test case 5: Recommends a Sleep Exercise                               | 41        |
| Test case 6: Improper Conversation Format                              | 42        |
| Test case 7: Different Language  | 43        |
| Test case 8: Text Data Augmentation                                    | 43        |
| Exercise function  | 44        |
| Decision table   | 45        |
| Test 4:  | 48        |
| Test 7:  | 49        |
| Augmented test case  | 57        |
| Invalid Topics Function  | 59        |
| Decision Table   | 59        |
| Test case 1:   | 60        |

|   |    |
|---|----|
| Test case 2:                                      | 60 |
| Test case 3:                                      | 61 |
| Test case 4:                                      | 61 |
| Test case 5:                                      | 62 |
| Test case 6:                                      | 62 |
| Test case 7:                                      | 63 |
| Test case 8:                                      | 63 |
| Test case 9:                                      | 64 |
| Test case 10:                                     | 64 |
| Section 3.2 Test case reports                     | 65 |
| Mental Illness Function                           | 65 |
| Test case 1: Recommends an exercise/technique     | 65 |
| Test case 2: Change in User's conversation format | 65 |
| Test case 3: Short Term Memory                    | 66 |
| Test case 4: Long Term Memory                     | 67 |
| Test case 5: Recommends Sleep Exercises           | 69 |
| Test case 6: Improper Conversation Format         | 70 |
| Test case 7: Different Language                   | 71 |
| Test case 8: Text Data Augmentation               | 72 |
| Test case 8.0: Normal Text                        | 72 |
| Test case 8.1: Random Word Swap                   | 73 |
| Test case 8.2: Random Word Delete                 | 73 |
| Test case 8.3: Random Word Insert                 | 74 |
| Test case 8.4: Synonym Augmentation               | 74 |
| Test case 8.5: OCR Augmentation                   | 75 |
| Test case 8.6: Keyboard Augmentation              | 76 |
| Test case 8.7: Random Char Insert                 | 76 |
| Test case 8.8: Random Char Swap                   | 77 |
| Test case 8.9: Random Char Delete                 | 78 |
| Exercise function                                 | 78 |
| Test case 1:                                      | 78 |
| Test case 2                                       | 79 |
| Test case 3                                       | 80 |
| Test case 4                                       | 80 |
| Test case 5                                       | 81 |
| Test case 6                                       | 82 |
| Test case 7                                       | 82 |
| Test case 8                                       | 83 |
| Test case 9                                       | 83 |
| Test case 10                                      | 84 |
| Test case 11                                      | 84 |

|   |           |
|---|-----------|
| Test case 12  | 85        |
| Test case 13  | 86        |
| Test case 14:   | 86        |
| Test case 15:   | 87        |
| Test case 16:   | 87        |
| Test case 17:   | 88        |
| Test case 18  | 88        |
| Augmented test case   | 89        |
| Invalid Topics Function   | 90        |
| Test case 1:  | 90        |
| Test case 2:  | 91        |
| Test case 3:  | 91        |
| Test case 4:  | 92        |
| Test case 5:  | 92        |
| Test case 6:  | 93        |
| Test case 7:  | 93        |
| Test case 8:  | 94        |
| Test case 9:  | 94        |
| Test case 10:   | 95        |
| Augmented test case   | 95        |
| Section 3.3 Test case analysis (statistics)   | 97        |
| Mental Illness Function   | 97        |
| Feeling Detection Function  | 97        |
| Exercise function   | 98        |
| Invalid Topics Function   | 99        |
| <b>Section 4 – AI Function Test Results and Test Criteria</b>   | <b>99</b> |
| Section 4.1 AI test model-based test results analysis test complexity, and test result statistics   | 99        |
| Mental Illness Function   | 99        |
| Feeling Detection Function  | 100       |
| Exercise function   | 102       |
| Invalid Topics Function   | 103       |
| Section 4.2 AI Function bug analysis  | 104       |
| Mental Illness Function   | 104       |
| Feeling Detection Function:   | 104       |
| Exercise function   | 105       |
| Invalid Topics Function   | 105       |
| Section 4.3 AI function test quality assessment   | 106       |
| Mental Illness Function   | 106       |
| • This test for detecting the User's mental illness and providing mental health support covers the app's capability to test the following | 106       |

|                         |     |
|-------------------------|-----|
| Exercise function       | 106 |
| Invalid Topics Function | 106 |

## Section 1 - AI Function Test Requirement Analysis

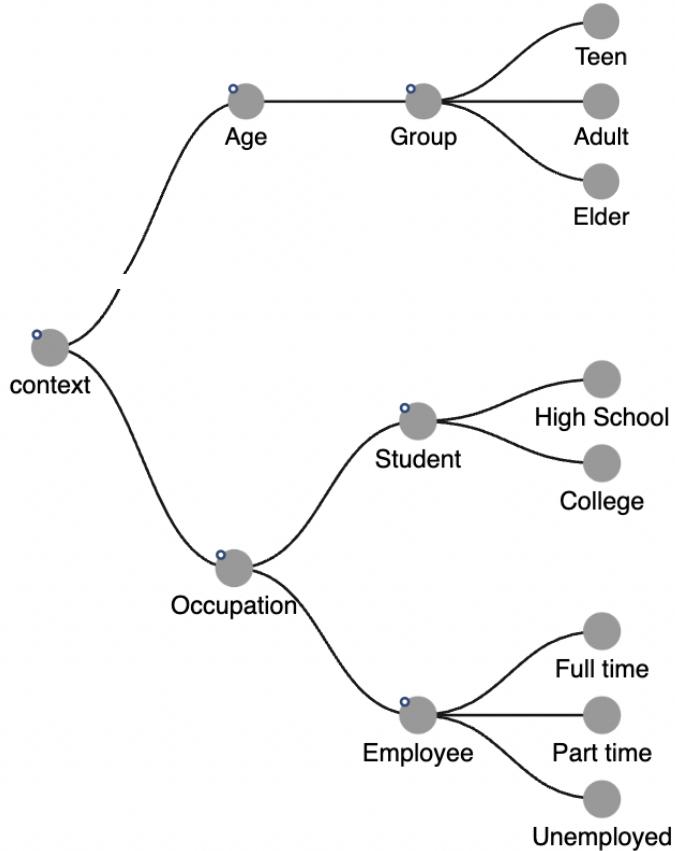
- We designed the test case based on AI test modeling
- Then we proceed to augmented the test case input
- We then test those test cases towards Wysa
- Wysa should pass if it behaves similar to the way we expected it to

## Section 2 - AI Test Modeling for your selected AI Features using 3D AI test tool (from Deliverable #2)

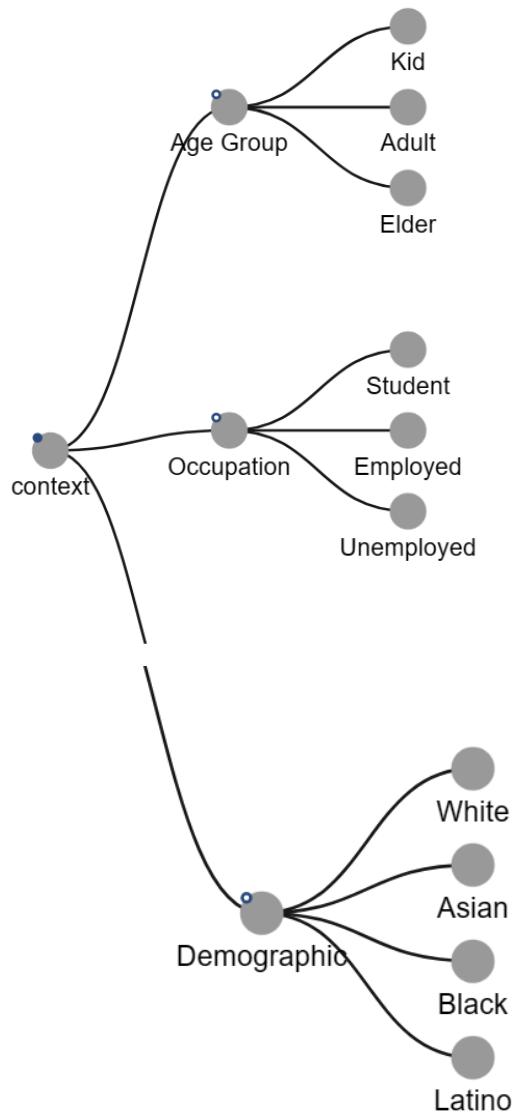
Some of the AI Test Modeling has been changed from Deliverable 2A

## Section 2.1 Context Modeling for AI Features

Mental Illness function



Exercise function



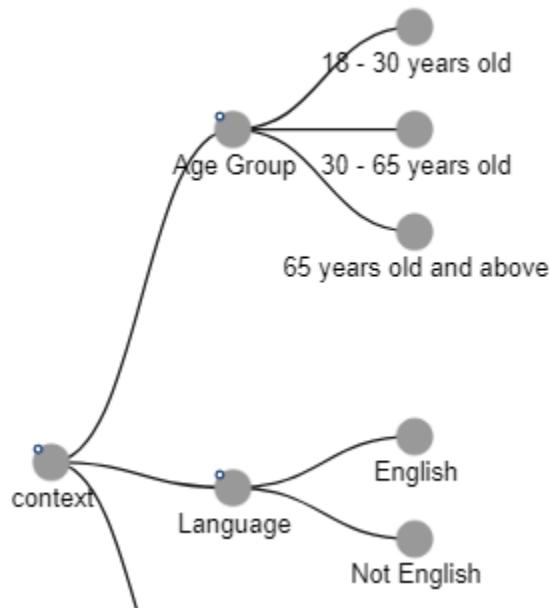
## Feeling Detection function

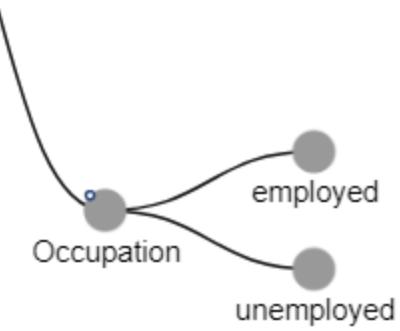
The context classification spanning tree for feeling detection encompasses four major cases and ten minor cases, providing a comprehensive framework to evaluate Wysa's performance in various contextual scenarios. The first major case focuses on age groups, encompassing three distinct age ranges: 18 to 30 years old, 30 to 65 years old, and 65+ years old. This allows for testing Wysa's functionality and effectiveness across different age groups, considering potential variations in emotional expression and needs across different age groups.

The second major case in the spanning tree is language, aiming to assess Wysa's ability to understand languages other than English. This case recognizes the importance of linguistic diversity and evaluates Wysa's performance in comprehending and responding to emotions expressed in languages other than English..

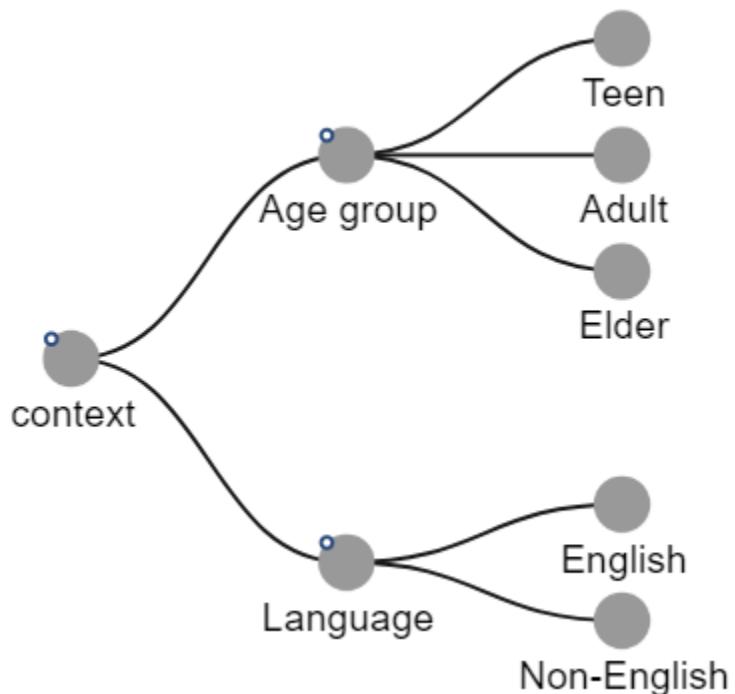
The last major case in the spanning tree is occupation, encompassing the user's occupation as a contextual factor. This case recognizes that an individual's occupation could impact their emotions, and evaluates Wysa's ability to understand and respond to emotions expressed in different professional contexts.

The seven minor cases within the major cases of age group, language, and occupation further provide an evaluation of Wysa's performance in various specific scenarios within each major case. This structured approach to testing allows for a comprehensive assessment of Wysa's capabilities in different contexts, providing insights into its effectiveness and potential areas for improvement in detecting and addressing emotions across diverse populations.



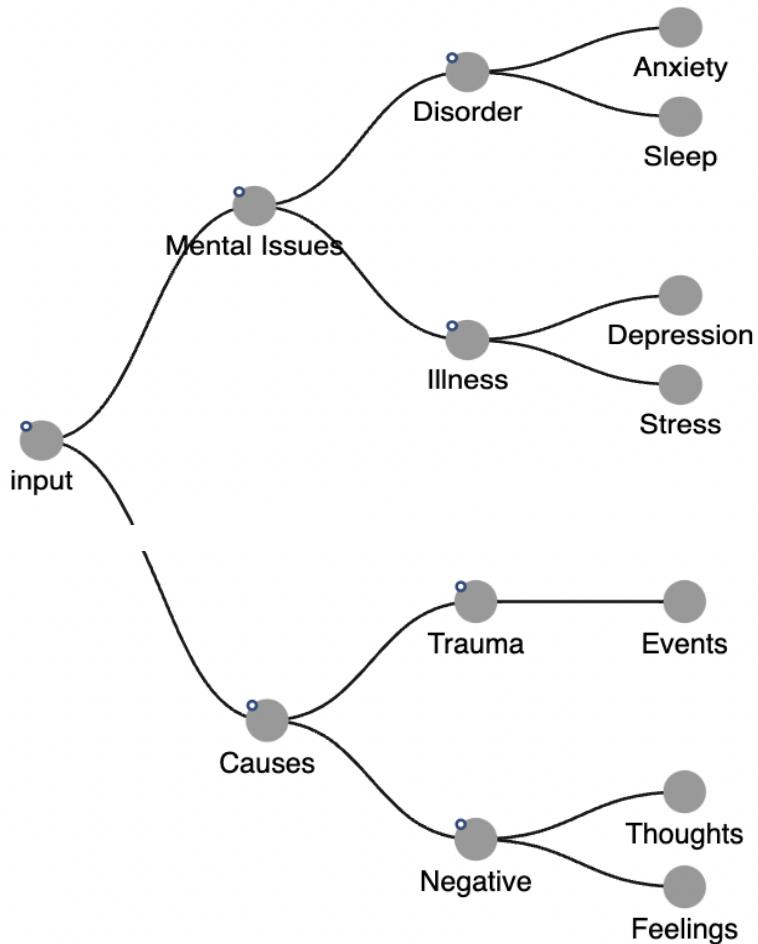


### Invalid Topics Function



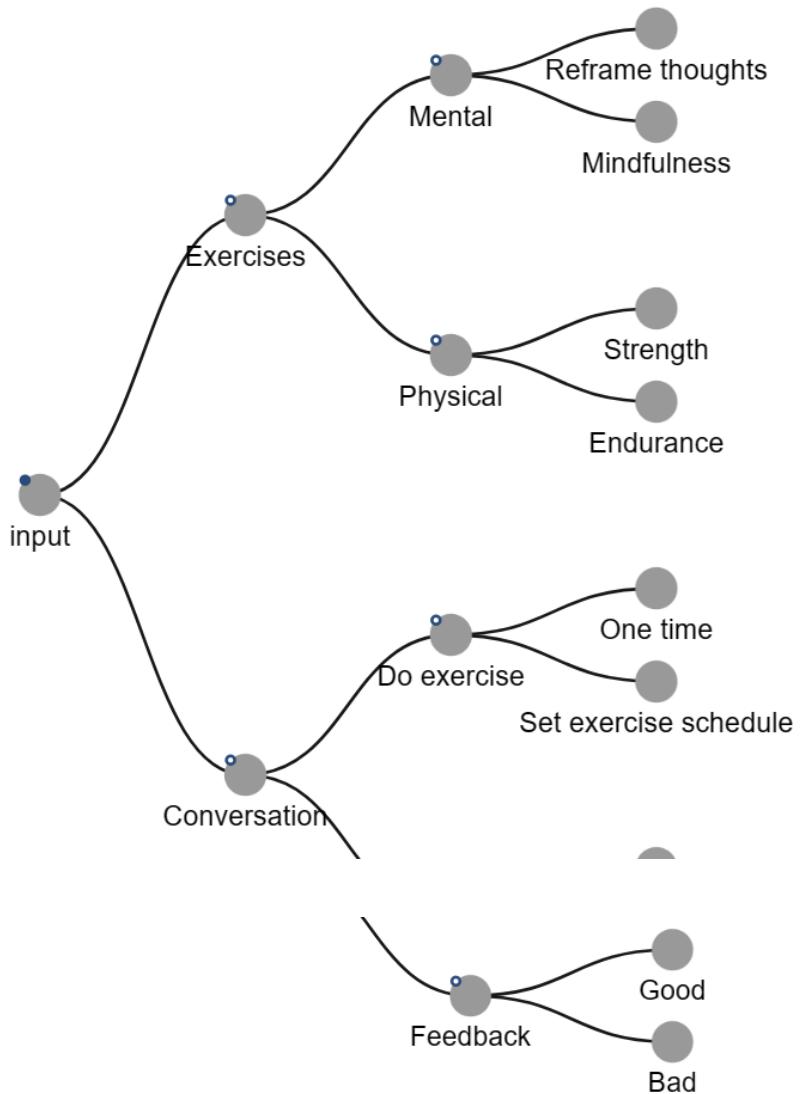
## Section 2.2 AI Function Input Classifications

Mental Illness function



Exercise function

Input tree

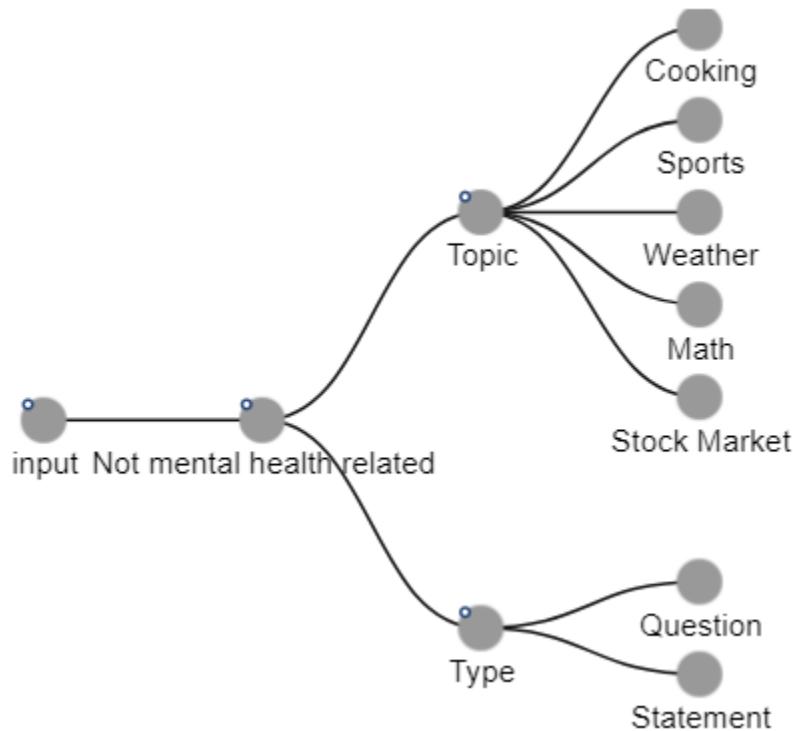


For our exercise input, we test 2 types of exercise, mental or physical.

For mental health we have the reframe thoughts exercise and mindfulness. Wysa has reframe thoughts functionality in order to reframe bad thoughts into positive ones, in order to help mental health problems. For physical exercise we have strength and endurance.

For conversation we have doing the exercise and feedback. For doing the exercise we have one time and then set an exercise schedule. For feedback we have good or bad feedback

## Invalid Topics Function



## Feeling Detection Function

### Input model

The input classification spanning tree for feeling detection encompasses a comprehensive framework that categorizes emotions into two major cases and five minor cases. The first major case is centered around negative or unpleasant feelings, which encompasses three minor cases including stress, sadness, and anxiety. These emotions are typically associated with distress, discomfort, or unease, and may manifest in various ways in different individuals.

The second major case in the classification spanning tree pertains to positive or pleasant feelings, including emotions such as happiness and excitement. These emotions are typically associated with joy, contentment, and enthusiasm, and are often characterized by a sense of well-being and positivity.

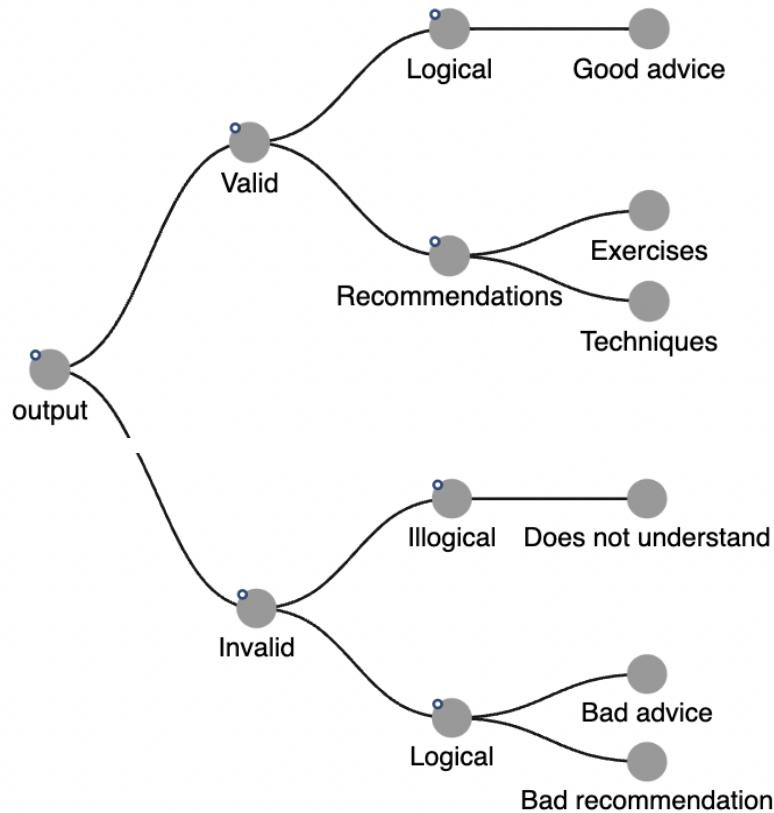
The aim of testing Wysa, the chatbot, was to evaluate its ability to accurately detect the user's emotions based on this classification spanning tree. By identifying and categorizing emotions into major and minor cases, Wysa aims to provide insights and support to users by understanding their emotional state and providing appropriate responses and guidance. Through this evaluation, it can be determined if Wysa is

effective in accurately detecting and responding to the user's emotions, thus offering valuable insights into the capabilities of the chatbot in the field of emotion detection and support.

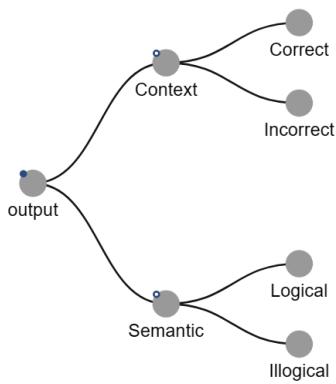


## Section 2.3 AI Function Output/Event/Action Classifications

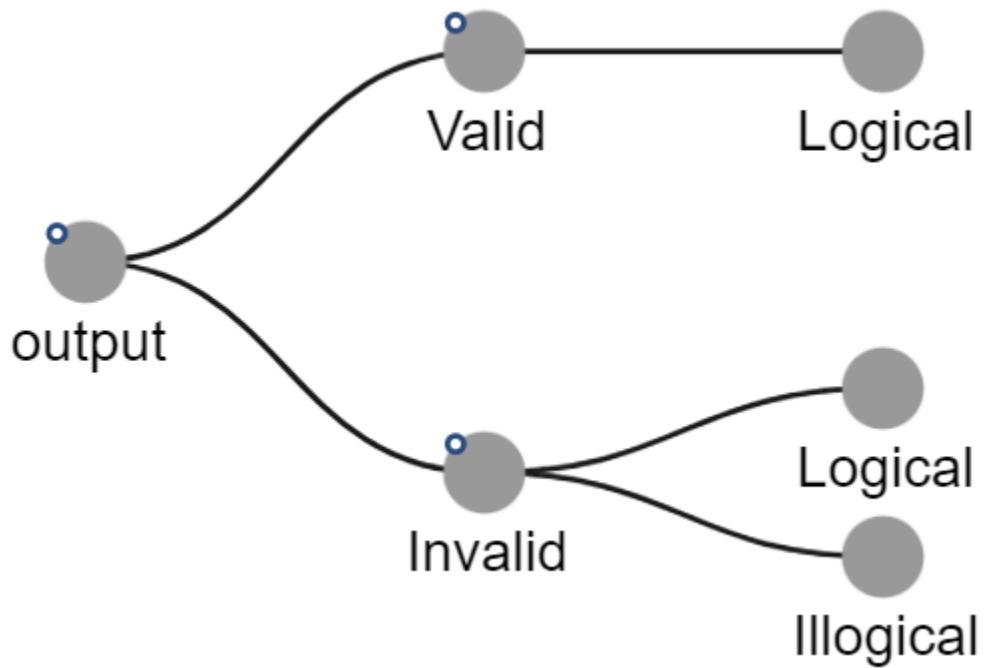
Mental Illness function



Exercise function



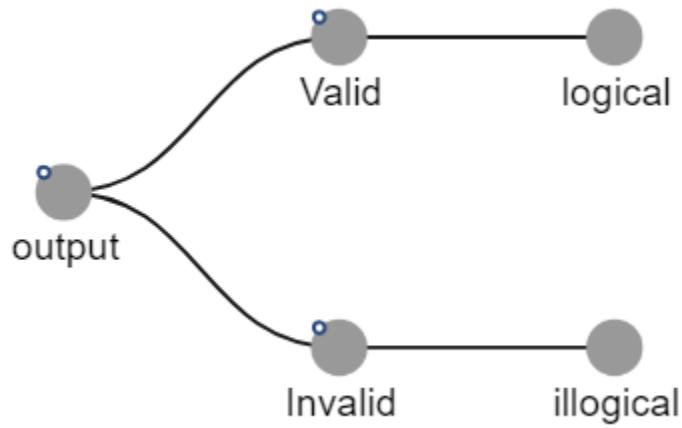
### Invalid Topics Function



### Feeling detection function

The output classification spanning tree for feeling detection comprises two major cases and two minor cases, providing a framework for evaluating the accuracy and logic of its responses. The first major case is labeled as "valid," which encompasses responses that are logical and coherent, aligning with the expected output from Wysa. These responses are deemed accurate and reliable, indicating that Wysa is providing meaningful and relevant answers to user queries.

The second major case in the classification tree is labeled as "invalid," encompassing responses that are illogical. The minor case of "illogical" includes responses that may not make sense or lack coherence, indicating that Wysa may not be providing logical answers.



## Section 2.4 AI Function Classification Decision Table (3D tables)

### Mental Illness function

| <b><i>id</i></b> | <b><i>Age.Group</i></b> | <b><i>Occupation.Employee</i></b> | <b><i>Occupation.Student</i></b> |
|------------------|-------------------------|-----------------------------------|----------------------------------|
| 1                | Teen                    | Full time                         | High School                      |
| 2                | Teen                    | Part time                         | High School                      |
| 3                | Teen                    | Unemployed                        | High School                      |
| 4                | Teen                    | Full time                         | College                          |
| 5                | Teen                    | Part time                         | College                          |
| 6                | Teen                    | Unemployed                        | College                          |
| 7                | Adult                   | Full time                         | High School                      |
| 8                | Adult                   | Part time                         | High School                      |
| 9                | Adult                   | Unemployed                        | High School                      |
| 10               | Adult                   | Full time                         | College                          |

| <b><i>id</i></b> | <b><i>Invalid.Ilogical</i></b> | <b><i>Invalid.Logical</i></b> | <b><i>Valid.Logical</i></b> | <b><i>Valid.Recommendations</i></b> |
|------------------|--------------------------------|-------------------------------|-----------------------------|-------------------------------------|
| 1                | undefined                      | undefined                     | undefined                   | undefined                           |
| 2                | undefined                      | undefined                     | undefined                   | undefined                           |
| 3                | undefined                      | undefined                     | undefined                   | undefined                           |
| 4                | undefined                      | undefined                     | undefined                   | undefined                           |
| 5                | undefined                      | undefined                     | undefined                   | undefined                           |
| 6                | undefined                      | undefined                     | undefined                   | undefined                           |
| 7                | undefined                      | undefined                     | undefined                   | undefined                           |
| 8                | undefined                      | undefined                     | undefined                   | undefined                           |

| <b><i>id</i></b> | <b><i>Invalid.Ilogical</i></b> | <b><i>Invalid.Logical</i></b> | <b><i>Valid.Logical</i></b> | <b><i>Valid.Recommendations</i></b> |
|------------------|--------------------------------|-------------------------------|-----------------------------|-------------------------------------|
| 1                | undefined                      | undefined                     | undefined                   | undefined                           |
| 2                | undefined                      | undefined                     | undefined                   | undefined                           |
| 3                | undefined                      | undefined                     | undefined                   | undefined                           |
| 4                | undefined                      | undefined                     | undefined                   | undefined                           |
| 5                | undefined                      | undefined                     | undefined                   | undefined                           |
| 6                | undefined                      | undefined                     | undefined                   | undefined                           |
| 7                | undefined                      | undefined                     | undefined                   | undefined                           |
| 8                | undefined                      | undefined                     | undefined                   | undefined                           |

| <b><i>id</i></b> | <b><i>Causes.Negative</i></b> | <b><i>Causes.Trauma</i></b> | <b><i>Mental.Issues.Disorder</i></b> | <b><i>Mental.Issues.Illness</i></b> |
|------------------|-------------------------------|-----------------------------|--------------------------------------|-------------------------------------|
| 1                | Thoughts                      | Events                      | Anxiety                              | Depression                          |
| 2                | Feelings                      | Events                      | Anxiety                              | Depression                          |
| 3                | Thoughts                      | Events                      | Anxiety                              | Stress                              |
| 4                | Feelings                      | Events                      | Sleep                                | Depression                          |
| 5                | Thoughts                      | Events                      | Sleep                                | Depression                          |
| 6                | Feelings                      | Events                      | Sleep                                | Stress                              |
| 7                | Thoughts                      | Events                      | Sleep                                | Stress                              |
| 8                | Feelings                      | Events                      | Sleep                                | Stress                              |

## Exercise function

| id | Conversation.Do exercise |      | Conversation.Feedback | Exercises.Mental | Exercises.Physical |
|----|--------------------------|------|-----------------------|------------------|--------------------|
|    | One time                 | Good | Reframe thoughts      | Strength         |                    |
| 1  | One time                 | Bad  | Mindfulness           | Strength         |                    |
| 2  | One time                 | Good | Mindfulness           | Strength         |                    |
| 3  | Set exercise schedule    | Bad  | Mindfulness           | Endurance        |                    |
| 4  | Set exercise schedule    | Good | Mindfulness           | Endurance        |                    |
| 5  | One time                 | Bad  | Mindfulness           | Endurance        |                    |
| 6  | One time                 | Good | Mindfulness           | Endurance        |                    |
| 7  | Set exercise schedule    | Bad  | Reframe thoughts      | Strength         |                    |
| 8  | Set exercise schedule    | Bad  | Reframe thoughts      | Strength         |                    |
| 9  | One time                 | Good |                       |                  |                    |
| 10 | Set exercise schedule    |      |                       |                  |                    |

| Id | Recommendations     |                 |               |                       | Feedback     |        | Action   |        | Status |          | Comments |          |
|----|---------------------|-----------------|---------------|-----------------------|--------------|--------|----------|--------|--------|----------|----------|----------|
|    | Invalid.Illogical   | Invalid.Logical | Valid.Logical | Valid.Recommendations | Conversation | Do     | Comments | Action | Status | Comments | Comments | Comments |
| 1  | Does not understand | Bad advice      | Good advice   | Exercises             | One time     | Guide  | Advice   | Start  | Open   | Good     | Advice   | Start    |
| 2  | undefined           | undefined       | Good advice   | Exercises             | One time     | Goal   | Advice   | Start  | Open   | Good     | Advice   | Start    |
| 3  | undefined           | undefined       | undefined     | undefined             | One time     | Bad    | Advice   | Start  | Open   | Bad      | Advice   | Start    |
| 4  | undefined           | undefined       | undefined     | undefined             | One time     | Goal   | Advice   | Start  | Open   | Good     | Advice   | Start    |
| 5  | undefined           | undefined       | undefined     | undefined             | One exercise | Advice | Advice   | Start  | Open   | Advice   | Advice   | Start    |
| 6  | undefined           | undefined       | undefined     | undefined             | One exercise | Goal   | Advice   | Start  | Open   | Advice   | Advice   | Start    |
| 7  | undefined           | undefined       | undefined     | undefined             | One exercise | Bad    | Advice   | Start  | Open   | Bad      | Advice   | Start    |
| 8  | undefined           | undefined       | undefined     | undefined             | One exercise | Goal   | Advice   | Start  | Open   | Good     | Advice   | Start    |
| 9  | undefined           | undefined       | undefined     | undefined             | One exercise | Bad    | Advice   | Start  | Open   | Bad      | Advice   | Start    |
| 10 | undefined           | undefined       | undefined     | undefined             | One exercise | Goal   | Advice   | Start  | Open   | Good     | Advice   | Start    |

| <i>id</i> | <i>Age Group</i> | <i>Demographic</i> | <i>Occupation</i> |
|-----------|------------------|--------------------|-------------------|
| 1         | Kid              | White              |                   |
| 2         | Kid              | Asian              | Student           |
| 3         | Kid              | Black              | Student           |
| 4         | Kid              | Latino             | Student           |
| 5         | Kid              | White              | Student           |
| 6         | Kid              | Asian              | Employed          |
| 7         | Kid              | Black              | Employed          |
| 8         | Kid              | Latino             | Employed          |
| 9         | Kid              | White              | Employed          |
| 10        | Kid              | Asian              | Unemployed        |
|           |                  |                    | Unemployed        |

## Invalid Topics Function

| id |           | Invalid | Valid     |    |                                 |                                |
|----|-----------|---------|-----------|----|---------------------------------|--------------------------------|
|    |           |         |           | id | Not mental health related.Topic | Not mental health related.Type |
| 1  | undefined |         | undefined | 1  | Cooking                         | Question                       |
| 2  | undefined |         | undefined | 2  | Stock Market                    | Statement                      |
| 3  | undefined |         | undefined | 3  | Cooking                         | Question                       |
| 4  | undefined |         | undefined | 4  | Sports                          | Statement                      |
| 5  | undefined |         | undefined | 5  | Sports                          | Question                       |
| 6  | undefined |         | undefined | 6  | Weather                         | Statement                      |
| 7  | undefined |         | undefined | 7  | Weather                         | Question                       |
| 8  | undefined |         | undefined | 8  | Math                            | Statement                      |
| 9  | undefined |         | undefined | 9  | Math                            | Question                       |
| 10 | undefined |         | undefined | 10 | Stock Market                    |                                |

| id | Age group | Language    |
|----|-----------|-------------|
| 1  | Teen      | English     |
| 2  | Teen      | Non-English |
| 3  | Adult     | English     |
| 4  | Adult     | Non-English |
| 5  | Elder     | English     |
| 6  | Elder     | Non-English |

Feeling detection function

3D decision table



Context model decision table

| ID | Age Group         | Language    | Occupation |
|----|-------------------|-------------|------------|
| C1 | 18 - 30 years old | -           | -          |
| C2 | -                 | Not English | -          |
| C3 | -                 | -           | Employed   |

Input model Decision table

| ID | Good Feelings | Bad Feelings |
|----|---------------|--------------|
| I1 | Happy         | -            |
| I2 | Excited       | -            |
| I3 | -             | Stress       |
| I4 | -             | Sad          |
| I5 | -             | Anxious      |

Output model decision table

| ID | Output    |
|----|-----------|
| O1 | Logical   |
| O2 | Illogical |

Test case design table

I have devised 8 test cases that encompass both the input and context model spanning trees. Each minor case will have a corresponding test case to ensure that Wysa accurately detects feelings, whether they are positive or negative, in various contextual scenarios. The first two test cases will focus on feeling detection with both invalid and valid conversation formats. The subsequent three test cases will solely test the accuracy of feeling detection. Lastly, the last three test cases will evaluate both feeling detection and contextual scenarios, such as age group, language, and occupation.

| Test Case ID | Input | Context | Output |
|--------------|-------|---------|--------|
| 1            | I1    | -       | O1     |
| 2            | I2    | -       | O2     |
| 3            | I3    | -       | O1     |
| 4            | I4    | -       | O1     |
| 5            | I5    | -       | O1     |

|   |    |    |    |
|---|----|----|----|
| 6 | I1 | C1 | O2 |
| 7 | I3 | C2 | O2 |
| 8 | I5 | C3 | O2 |

## Section 3 – AI Function Test Cases with Inputs/Expected Outputs

### Section 3.1 Test data models

Feeling Detection function

|                 |  |
|-----------------|--|
| Test case ID    | 1  |
| Description     | Happiness detection with valid conversation format   |
| Purpose         | To test Wysa to see if Wysa could detect happiness   |
| Input           | "I'm feeling so happy today!"  |
| Output          | "I'd love to hear about what makes you feel happy today"<br>"What would make today better" |
| Expected Output | "That's great to hear! What made you happy today?"   |
| Result          | Passed   |

|                 |  |
|-----------------|--|
| Test case ID    | 2  |
| Description     | excitement detection with invalid conversation format        |
| Purpose         | To test Wysa to see if Wysa could detect excitement          |
| Input           | "I won the lottery today!"<br>"I just got promoted at work!" |
| Output          | "What would you require to focus on that?"                   |
| Expected Output | "I'm having trouble understanding your message. Can you      |

|        |  |
|--------|--|
|        | please provide more details or follow a proper conversation format?" |
| Result | failed   |

|                 |   |
|-----------------|---|
| Test case ID    | 3   |
| Description     | Stress detection  |
| Purpose         | To test Wysa to see if Wysa could detect stress   |
| Input           | "Work has been really overwhelming lately, and I'm feeling stressed."   |
| Output          | "That's an uncomfortable feeling, I know. Observe how it feels for a bit, just allowing it to exist."                                     |
| Expected Output | "It sounds like you're going through a stressful time with your exams. Remember to take breaks and practice self-care during this period" |
| Result          | Passed  |

|                 |  |
|-----------------|--|
| Test case ID    | 4  |
| Description     | Sad detection  |
| Purpose         | To test Wysa to see if Wysa could detect sadness   |
| Input           | "I'm feeling really sad today"   |
| Output          | "I understand you're feeling sad. It's only natural to feel this way"  |
| Expected Output | "I'm sorry to hear that you're feeling sad. Remember to take care of yourself and reach out to supportive people around you" |
| Result          | Passed   |

|                 |  |
|-----------------|--|
| Test case ID    | 5  |
| Description     | Anxiety detection  |
| Purpose         | To test Wysa to see if Wysa could detect anxiety   |
| Input           | "I'm feeling really anxious about my presentation tomorrow"  |
| Output          | "I'm here to help you feel more in control"  |
| Expected Output | "It sounds like you're feeling anxious about your presentation. Remember to practice deep breathing and positive self-talk to manage your anxiety" |
| Result          | Passed   |

|                 |   |
|-----------------|---|
| Test case ID    | 6   |
| Description     | Happy detection with context of a person who is from 18 - 30 years old  |
| Purpose         | To test Wysa to see if Wysa could detect happiness expressed by a user who is from 18 to 30 years old   |
| Input           | "I aced my exam and I'm so happy!"  |
| Output          | "Scientists say that positive emotions increase one's awareness and encourage novel thoughts and actions. This could be a fantastic day to start something creative." |
| Expected Output | "Congratulations on acing your exam! That's awesome! How do you plan to celebrate?"   |
| Result          | Failed  |

|              |  |
|--------------|--|
| Test case ID | 7  |
| Description  | Stress detection in non English language   |
| Purpose      | To test Wysa to see if Wysa could detect stress in non English language                |
| Input        | "Estoy muy estresado por los exámenes" (Spanish for "I am very stressed about exams"). |

|                 |   |
|-----------------|---|
| Output          | "That looks like a bunch of letters that don't make a lot of sense together"<br>"Are you testing my AI capabilities?"   |
| Expected Output | "Comprendo que te sientas estresado(a) por los exámenes. Trata de descansar y respirar profundamente" (Spanish for "I understand that you feel stressed about exams. Try to rest and breathe deeply") |
| Result          | Failed  |

|                 |   |
|-----------------|---|
| Test case ID    | 8   |
| Description     | Anxiety for a employed user   |
| Purpose         | To test Wysa to see if Wysa could detect anxiety expressed by a employed user   |
| Input           | "I'm feeling really anxious about the upcoming presentation at work"  |
| Output          | "This kind of distress can feel overwhelming. I can help you learn ways to feel more in control"                            |
| Expected Output | "It's understandable to feel anxious about the upcoming presentation. Remember to take deep breaths and practice self-care" |
| Result          | Failed  |

### Data Augmentation for Feeling detection function

This section involves conducting tests on eight previously identified test cases using data augmentation techniques. The aim is to check Wysa's capability to identify emotions in input text, even with certain variations in the input caused by data augmentation. The three types of data augmentation techniques employed are random word swap, random word insertion, and random word deletion. Through this analysis, we can evaluate Wysa's performance in identifying emotions and its robustness to handle variations in input text.

|                     |   |
|---------------------|---|
| <b>Test Case ID</b> | <b>1.1</b>  |
| Description         | excitement detection with invalid conversation format |
| Data Augmentation   | Random word swap                                      |
| Input               | "I'm happy feeling today!"                            |
| Result              | Passed  |
| <b>Test Case ID</b> | <b>1.2</b>  |
| Data Augmentation   | Random word delete                                    |
| Input               | "feeling so today!"                                   |
| Result              | Failed  |
| <b>Test Case ID</b> | <b>1.3</b>  |
| Data Augmentation   | Random word insert                                    |
| Input               | "I'm feeling so glad happy today!"                    |
| Result              | Failed  |
| Expected output     | "That's great to hear! What made you happy today?"    |

|                     |  |
|---------------------|--|
| <b>Test Case ID</b> | <b>2.1</b>   |
| Description         | Happiness detection with valid conversation format |
| Data Augmentation   | Random word swap                                   |
| Input               | "I today! the lottery won"                         |
| Result              | failed   |
| <b>Test Case ID</b> | <b>2.2</b>   |

|                     |  |
|---------------------|--|
| Data Augmentation   | Random word delete   |
| Input               | "won the today!"   |
| Result              | failed   |
| <b>Test Case ID</b> | <b>2.3</b>   |
| Data Augmentation   | Random word insert   |
| Input               | "I won win the lottery today!"   |
| Result              | failed   |
| Expected output     | "I'm having trouble understanding your message. Can you please provide more details or follow a proper conversation format?" |

|                     |  |
|---------------------|--|
| <b>Test Case ID</b> | <b>3.1</b>   |
| Description         | Stress detection   |
| Data Augmentation   | Random word swap   |
| Input               | "Work has stressed. really overwhelming lately, and I'm feeling been"      |
| Result              | failed   |
| <b>Test Case ID</b> | <b>3.2</b>   |
| Data Augmentation   | Random word delete   |
| Input               | "has been overwhelming lately, and I'm"                                    |
| Result              | failed   |
| <b>Test Case ID</b> | <b>3.3</b>   |
| Data Augmentation   | Random word insert   |
| Input               | "Work has been really overwhelming lately, and I'm feeling real stressed." |

|                 |   |
|-----------------|---|
| Result          | failed  |
| Expected output | "It sounds like you're going through a stressful time with your exams. Remember to take breaks and practice self-care during this period" |

|                     |  |
|---------------------|--|
| <b>Test Case ID</b> | <b>4.1</b>   |
| Description         | Sad detection  |
| Data Augmentation   | Random word swap   |
| Input               | "I'm today really sad feeling"   |
| Result              | Passed   |
| <b>Test Case ID</b> | <b>4.2</b>   |
| Data Augmentation   | Random word delete   |
| Input               | "feeling really today"   |
| Result              | Failed   |
| <b>Test Case ID</b> | <b>4.3</b>   |
| Data Augmentation   | Random word insert   |
| Input               | "I'm feeling really sad intuitive feeling today"   |
| Result              | Passed   |
| Expected output     | "I'm sorry to hear that you're feeling sad. Remember to take care of yourself and reach out to supportive people around you" |

|                     |                   |
|---------------------|-------------------|
| <b>Test Case ID</b> | <b>5.1</b>        |
| Description         | Anxiety detection |

|                     |  |
|---------------------|--|
| Data Augmentation   | Random word swap   |
| Input               | "I'm really feeling anxious about my presentation tomorrow"  |
| Result              | Passed   |
| <b>Test Case ID</b> | <b>5.2</b>   |
| Data Augmentation   | Random word delete   |
| Input               | "feeling really about my presentation tomorrow"  |
| Result              | Failed   |
| <b>Test Case ID</b> | <b>5.3</b>   |
| Data Augmentation   | Random word insert   |
| Input               | "I'm intuitive feeling feeling really anxious about my presentation tomorrow"  |
| Result              | failed   |
| Expected output     | "It sounds like you're feeling anxious about your presentation. Remember to practice deep breathing and positive self-talk to manage your anxiety" |

|                     |  |
|---------------------|--|
| <b>Test Case ID</b> | <b>6.1</b>   |
| Description         | Happy detection with context of a person who is from 18 - 30 years old |
| Data Augmentation   | Random word swap   |
| Input               | "I my aced exam and I'm so happy!"                                     |
| Result              | Passed   |
| <b>Test Case ID</b> | <b>6.2</b>   |
| Data Augmentation   | Random word delete   |

|                     |   |
|---------------------|---|
| Input               | "aced my and I'm so happy!"   |
| Result              | Passed  |
| <b>Test Case ID</b> | <b>6.3</b>  |
| Data Augmentation   | Random word insert  |
| Input               | "I sail through aced my exam and I'm so happy!"                                     |
| Result              | Passed  |
| Expected output     | "Congratulations on acing your exam! That's awesome! How do you plan to celebrate?" |

|                     |  |
|---------------------|--|
| <b>Test Case ID</b> | <b>7.1</b>   |
| Description         | Stress detection in non English language                                       |
| Data Augmentation   | Random word swap   |
| Input               | "Estoy los estresado por muy exámenes"   |
| Result              | Failed   |
| <b>Test Case ID</b> | <b>7.2</b>   |
| Data Augmentation   | Random word delete   |
| Input               | "muy estresado los exámenes"   |
| Result              | Failed   |
| <b>Test Case ID</b> | <b>7.3</b>   |
| Data Augmentation   | Random word insert   |
| Input               | "Estoy muy estresado por los exámenes"   |
| Result              | Failed   |
| Expected output     | "Comprendo que te sientas estresado(a) por los exámenes. Trata de descansar y" |

|  |   |
|--|---|
|  | respirar profundamente" (Spanish for "I understand that you feel stressed about exams. Try to rest and breathe deeply") |
|--|---|

|                     |   |
|---------------------|---|
| <b>Test Case ID</b> | <b>8.1</b>  |
| Description         | Anxiety detection for an employed user  |
| Data Augmentation   | Random word swap  |
| Input               | "I'm feeling work anxious about the upcoming presentation at really"  |
| Result              | Failed  |
| <b>Test Case ID</b> | <b>8.2</b>  |
| Data Augmentation   | Random word delete  |
| Input               | "feeling really about the upcoming presentation"  |
| Result              | Failed  |
| <b>Test Case ID</b> | <b>8.3</b>  |
| Data Augmentation   | Random word insert  |
| Input               | " I'm feeling really anxious about the upcoming presentation at intuitive feeling work"                                     |
| Result              | Failed  |
| Expected output     | "It's understandable to feel anxious about the upcoming presentation. Remember to take deep breaths and practice self-care" |

## Mental Illness function

|    |                 |               | Mental<br>Issues.Disorder | Mental<br>Issues.Illness |
|----|-----------------|---------------|---------------------------|--------------------------|
| id | Causes.Negative | Causes.Trauma |                           |                          |
| 1  | Thoughts        | Events        | Anxiety                   | Depression               |
| 2  | Feelings        | Events        | Anxiety                   | Depression               |
| 3  | Thoughts        | Events        | Anxiety                   | Stress                   |
| 4  | Feelings        | Events        | Anxiety                   | Stress                   |
| 5  | Thoughts        | Events        | Sleep                     | Depression               |
| 6  | Feelings        | Events        | Sleep                     | Depression               |
| 7  | Thoughts        | Events        | Sleep                     | Stress                   |
| 8  | Feelings        | Events        | Sleep                     | Stress                   |

Decision Table

|   | Causes   |          |        | Mental Illnesses |       |            |        |
|---|----------|----------|--------|------------------|-------|------------|--------|
|   | Negative |          | Trauma | Disorder         |       | Illness    |        |
|   | Thoughts | Feelings | Events | Anxiety          | Sleep | Depression | Stress |
| 1 | X        |          | X      | X                |       | X          |        |
| 2 |          | X        | X      | X                |       | X          |        |
| 3 | X        |          | X      | X                |       |            | X      |
| 4 |          | X        | X      | X                |       |            | X      |
| 5 | X        |          | X      |                  | X     | X          |        |
| 6 |          | X        | X      |                  | X     | X          |        |
| 7 | X        |          | X      |                  | X     |            | X      |

|   |  |   |   |  |   |  |   |
|---|--|---|---|--|---|--|---|
| 8 |  | X | X |  | X |  | X |
|---|--|---|---|--|---|--|---|

Test case 1: Recommends an exercise/technique

| Test Case ID          | M1  |               |                        |   |  |
|-----------------------|---|---------------|------------------------|---|--|
| Test Case Description | User has anxiety and depression that is caused by negative thoughts and traumatic events. |               |                        |   |  |
| Input                 | Causes.Negative   | Causes.Trauma | Mental Issues.Disorder | Mental Issues.Illness                     |  |
|                       | Thoughts  | Events        | Anxiety                | Depression                                |  |
| Steps                 | Wysa  |               | User                   |   |  |
| 1                     | Asks User how they are feeling  |               |                        |   |  |
| 2                     |   |               |                        | Says, "I'm feeling depressed"             |  |
| 3                     | Asks User why they are depressed  |               |                        |   |  |
| 4                     |   |               |                        | Says, "I have anxiety"                    |  |
| 5                     | Asks User to share one thought about the situation  |               |                        |   |  |
| 6                     |   |               |                        | States a thought                          |  |
| 7                     | Lists different thoughts that might cause the way the User is feeling                     |               |                        |   |  |
| 8                     | Asks User which thought applies to them   |               |                        |   |  |
| 9                     |   |               |                        | Chooses which thought they relate to most |  |
| 10                    | Guides User on ways to make them feel better based on the User's thought                  |               |                        |   |  |
| 11                    |   |               |                        | Follows guide                             |  |
| Expected              | Wysa recommends an exercise/technique to help the User feel better.                       |               |                        |   |  |

|        |  |
|--------|--|
| Output |  |
|--------|--|

Test case 2: Change in User's conversation format

|                       |   |               |  |                       |
|-----------------------|---|---------------|--|-----------------------|
| Test Case ID          | M2  |               |  |                       |
| Test Case Description | User has anxiety and depression that is caused by negative feelings and traumatic events.               |               |  |                       |
| Input                 | Causes.Negative   | Causes.Trauma | Mental Issues.Disorder   | Mental Issues.Illness |
|                       | Feelings  | Events        | Anxiety  | Depression            |
| Steps                 | Wysa  |               | User   |                       |
| 1                     | Asks how the User is feeling  |               |  |                       |
| 2                     |   |               | Says, "What are different exercises that will help with my anxiety?" |                       |
| 3                     | Asks what the cause of their anxiety  |               |  |                       |
| 4                     |   |               | States that negative feelings is the cause of their anxiety          |                       |
| 5                     | Provides a list of types of feelings that the user might be going through                               |               |  |                       |
| 6                     |   |               | Chooses the feeling that they relate to most                         |                       |
| 7                     | Asks to confirm if this feeling is what they are feeling  |               |  |                       |
| 8                     |   |               | Says, "Let's do something else"                                      |                       |
| 9                     | Asks User what they want to do  |               |  |                       |
| Expected Output       | Wysa should be able to adjust to the User's conversation format and respond to the User's mental issue. |               |  |                       |

### Test case 3: Short Term Memory

|                       |  |               |  |                       |
|-----------------------|--|---------------|--|-----------------------|
| Test Case ID          | M3   |               |  |                       |
| Test Case Description | User is a student that is feeling stressed because of too much homework and creates a positive thought                   |               |  |                       |
| Input                 | Causes.Negative  | Causes.Trauma | Mental Issues.Disorder                   | Mental Issues.Illness |
|                       | Thoughts   | Events        | Anxiety                                  | Stress                |
| Steps                 | Wysa   |               | User                                     |                       |
| 1                     | Asks User how they are feeling   |               |  |                       |
| 2                     |  |               | Says "I'm feeling stressed"              |                       |
| 3                     | Asks if User is also anxious   |               |  |                       |
| 4                     |  |               | Clicks "Yes"                             |                       |
| 5                     | Asks if it is affecting the User physically too  |               |  |                       |
| 6                     | Prompts User to choose, "Mostly mind", "Mostly physical", or "A bit of both"   |               |  |                       |
| 7                     |  |               | User chooses "Mostly mind"               |                       |
| 8                     | Asks User what thought comes to mind that causes their stress  |               |  |                       |
| 9                     |  |               | User says "I have a lot of homework"     |                       |
| 10                    | Asks the user what change they would hope for  |               |  |                       |
| 11                    |  |               | User says what they would like to change |                       |
| 12                    | Restates the user's thought, "I have a lot of homework" and asks what positives about the situation/themselves have been |               |  |                       |

|                 |  |                               |
|-----------------|--|-------------------------------|
|                 | ignored  |                               |
| 13              |  | User creates positive thought |
| Expected Output | Wysa should be able to remember User's mental health and thought input to help User feel better. |                               |

#### Test case 4: Long Term Memory

|                       |  |               |                                    |                       |
|-----------------------|--|---------------|------------------------------------|-----------------------|
| Test Case ID          | M4   |               |                                    |                       |
| Test Case Description | User is a student that is feeling stressed because of a midterm and starts a new conversation after creating a positive thought. |               |                                    |                       |
| Input                 | Causes.Negative  | Causes.Trauma | Mental Issues.Disorder             | Mental Issues.Illness |
|                       | Feelings   | Events        | Anxiety                            | Stress                |
| Steps                 | Wysa   |               | User                               |                       |
| 1                     | Asks User how they are feeling   |               |                                    |                       |
| 2                     |  |               | Says "I'm feeling stressed"        |                       |
| 3                     | Asks if User is also anxious   |               |                                    |                       |
| 4                     |  |               | Clicks "Yes"                       |                       |
| 5                     | Asks if it is affecting the User physically too  |               |                                    |                       |
| 6                     | Prompts User to choose, "Mostly mind", "Mostly physical", or "A bit of both"   |               |                                    |                       |
| 7                     |  |               | User chooses "Mostly mind"         |                       |
| 8                     | Asks User what thought comes to mind that causes their stress  |               |                                    |                       |
| 9                     |  |               | User says "I have a midterm today" |                       |
| 10                    | Asks the user what change they   |               |                                    |                       |

|                 |  |  |
|-----------------|--|--|
|                 | would hope for   |  |
| 11              |  | User says what they would like to change |
| 12              | Restates the user's thought, "I have a lot of homework" and asks what positives about the situation/themselves have been ignored |  |
| 13              |  | User creates positive thought            |
| 14              | Prompts the User to "Start new Chat"   |  |
| 15              |  | User presses "Start new Chat"            |
| 16              | Wysa states that User has mentioned that they're "feeling stressed"  |  |
| Expected Output | Wysa should be able to remember the User's feeling of stress and event that causes it even after the User creates a new chat     |  |

#### Test case 5: Recommends a Sleep Exercise

|                       |   |               |                               |                       |
|-----------------------|---|---------------|-------------------------------|-----------------------|
| Test Case ID          | M5  |               |                               |                       |
| Test Case Description | User needs help to sleep.                       |               |                               |                       |
| Input                 | Causes.Negative                                 | Causes.Trauma | Mental Issues.Disorder        | Mental Issues.Illness |
|                       | Thoughts  | Events        | Sleep                         | Depression            |
| Steps                 | Wysa  |               | User                          |                       |
| 1                     |   |               | Says, "I am lacking in sleep" |                       |
| 2                     | Asks User if they would like to help them sleep |               |                               |                       |
| 3                     |   |               | Clicks, "Yes"                 |                       |

|                 |   |  |
|-----------------|---|--|
| 4               | Asks if they are also feeling tired   |  |
| 5               |   | Can choose, "Yes" or "Not really"                  |
| 6               | Asks User if they would like to try a meditation to help them sleep                               |  |
| 7               |   | Clicks, "No"                                       |
| 8               | Provides all sleep meditations and relaxation exercises   |  |
| 9               |   | User chooses a sleep recommendation and follows it |
| Expected Output | Wysa should be able to provide sleep meditations and relaxation exercises to help the User sleep. |  |

#### Test case 6: Improper Conversation Format

|                       |  |               |   |                       |
|-----------------------|--|---------------|---|-----------------------|
| Test Case ID          | M6   |               |   |                       |
| Test Case Description | User needs help staying awake and wants to know why they are depressed and having trouble sleeping |               |   |                       |
| Input                 | Causes.Negative  | Causes.Trauma | Mental Issues.Disorder                    | Mental Issues.Illness |
|                       | Feelings   | Events        | Sleep                                     | Depression            |
| Steps                 | Wysa   |               | User                                      |                       |
| 1                     |  |               | Asks, "Why am I having trouble sleeping?" |                       |
| 2                     | Asks what kind of sleeping problems the User has   |               |   |                       |
| 3                     |  |               | Asks, "Can you help me stay up?"          |                       |
| 4                     | Provides physical activities that will help the User stay awake                                    |               |   |                       |
| 5                     |  |               | Asks, "Why am I depressed?"               |                       |
| 6                     | Asks what happened that made   |               |   |                       |

|                 |  |  |
|-----------------|--|--|
|                 | the user feel depressed  |  |
| Expected Output | Wysa should be able to provide the right recommendations to the user and ask questions that will help determine the User's mental condition. |  |

#### Test case 7: Different Language

|                       |   |               |                        |                       |
|-----------------------|---|---------------|------------------------|-----------------------|
| Test Case ID          | M7  |               |                        |                       |
| Test Case Description | User tells Wysa their mental illness in a different language.   |               |                        |                       |
| Input                 | Causes.Negative   | Causes.Trauma | Mental Issues.Disorder | Mental Issues.Illness |
|                       | Thoughts  | Events        | Sleep                  | Stress                |
| Steps                 | Wysa  |               | User                   |                       |
| 1                     | Asks how the User is feeling  |               |                        |                       |
| 2                     |   |               | Says, “ストレスがたまります”     |                       |
| 3                     | Asks why the user is feeling stressed   |               |                        |                       |
| 4                     |   |               | Says, “仕事は疲れる”         |                       |
| 5                     | Asks what thought the User is having that causes the stress   |               |                        |                       |
| Expected Output       | Wysa should be able to understand the User in a different language and ask questions regarding their condition. |               |                        |                       |

#### Test case 8: Text Data Augmentation

|                       |   |              |        |        |
|-----------------------|---|--------------|--------|--------|
| Test Case ID          | M8  |              |        |        |
| Test Case Description | User states that they are stressed out. Using the AI tool, we will be able to input different augmented text inputs for Step 2. |              |        |        |
| Input                 | Causes.Negative   | Causes.Traum | Mental | Mental |

|                 |   | a      | Issues.Disorder                            | Issues.Illness |
|-----------------|---|--------|--|----------------|
|                 | Feelings  | Events | Sleep                                      | Stress         |
| Steps           | Wysa  |        | User                                       |                |
| 1               | Asks how is the User's day  |        |  |                |
| 2               |   |        | User responds "I am feeling stressed out." |                |
| 3               | States that this feeling brings a lot of unpleasantness and advises User to take this as a positive push. |        |  |                |
| 4               | Asks User why they are feeling stressed   |        |  |                |
| Expected Output | Wysa should be able to recognize that the User is stressed.   |        |  |                |

## Exercise function

New exercise function test models been made and as below

| id | Conversation.Do exercise | Conversation.Feedback | Exercises.Mental | Exercises.Physical   |
|----|--------------------------|-----------------------|------------------|----------------------|
| 1  | One time                 | Good                  | Reframe thoughts | Target muscle groups |
| 2  | One time                 | Bad                   | Mindfulness      | Target muscle groups |
| 3  | Set exercise schedule    | Good                  | Mindfulness      | Target muscle groups |
| 4  | Set exercise schedule    | Bad                   | Mindfulness      | Full body workout    |
| 5  | One time                 | Good                  | Mindfulness      | Full body workout    |
| 6  | One time                 | Bad                   | Mindfulness      | Full body workout    |
| 7  | Set exercise schedule    | Good                  | Mindfulness      | Full body workout    |
| 8  | Set exercise schedule    | Bad                   | Mindfulness      | Full body workout    |
| 9  | One time                 | Bad                   | Reframe thoughts | Target muscle groups |
| 10 | Set exercise schedule    | Good                  | Reframe thoughts | Target muscle groups |

New decision table because the one above generated by AI testing is overflow

## Decision table

|  | Exercises   |                 |          |            | Conversation |              |          |     |
|--|-------------|-----------------|----------|------------|--------------|--------------|----------|-----|
|  | Mental      |                 | Physical |            | Exercises    |              | Feedback |     |
|  | Mindfulness | Reframe thought | Strength | Durability | Do one time  | Set schedule | Good     | Bad |

|    |   |   |   |   |   |   |   |   |
|----|---|---|---|---|---|---|---|---|
|    |   | s |   |   |   |   |   |   |
| 1  | X |   |   |   | X |   |   |   |
| 2  | X |   |   |   |   | X |   |   |
| 3  | X |   |   |   |   |   | X |   |
| 4  | X |   |   |   |   |   |   | X |
| 5  |   | X |   |   | X |   |   |   |
| 6  |   | X |   |   |   | X |   |   |
| 7  |   | X |   |   |   |   | X |   |
| 8  |   | X |   |   |   |   |   | X |
| 9  |   |   | X |   | X |   |   |   |
| 10 |   |   | X |   |   | X |   |   |
| 11 |   |   | X |   |   |   | X |   |
| 12 |   |   | X |   |   |   |   | X |
| 13 |   |   |   | X | X |   |   |   |
| 14 |   |   |   | X |   | X |   |   |
| 15 |   |   |   | X |   |   | X |   |
| 16 |   |   |   | X |   |   |   | X |

Test 1:

|                       |   |   |  |
|-----------------------|---|---|--|
| Test case ID          | E1  |   |  |
| Test case description | User wants to do mindfulness exercises (once) |   |  |
| Steps                 | Wysa  | User  |  |
| 1                     |   | "Give me some mindfulness exercises for Asian elders" |  |
| 2                     | Wysa offers some specialized mindfulness      |   |  |

|                 |   |                                    |
|-----------------|---|------------------------------------|
|                 | exercises   |                                    |
| 3               |   | User chooses one of the exercises  |
| 4               | Wysa guides user through the exercise   | User follows the exercises         |
| 5               |   | User asks what exercises been done |
| Expected output | <ul style="list-style-type: none"> <li>• Wysa should be able to guide user through exercises and remember what exercises been done</li> </ul> |                                    |

Test 2:

|                       |  |   |
|-----------------------|--|---|
| Test case ID          | E2   |   |
| Test case description | User wants to set exercise routines for mindfulness exercises  |   |
| Steps                 | Wysa   | User  |
| 1                     |  | “I want to set some exercise routines as an Latino adult” |
| 2                     | Wysa guides user the exercise routine process  | User follows the exercise routine process                 |
| 3                     |  | User asks what exercises routines been set up             |
| Expected output       | <ul style="list-style-type: none"> <li>• Wysa should be able to guide user through exercise routines and remember the routines set up</li> </ul> |   |

Test 3:

|                       |   |   |
|-----------------------|---|---|
| Test case ID          | E3  |   |
| Test case description | User wants to give good feedback to mindfulness exercises |   |
| Steps                 | Wysa  | User  |
| 1                     |   | “Give me some mindfulness exercises for Asian elders” |

|                 |  |                                    |
|-----------------|--|------------------------------------|
| 2               | Wysa offers some personalized mindfulness exercises  |                                    |
| 3               |  | User chooses one of the exercises  |
| 4               | Wysa guides user through the exercise  | User follows the exercises         |
| 5               |  | "This exercise was helpful for me" |
| Expected output | <ul style="list-style-type: none"> <li>Wysa should be able to guide user through exercises and incorporate the feedback</li> </ul> |                                    |

Test 4:

|                       |  |   |
|-----------------------|--|---|
| Test case ID          | E4   |   |
| Test case description | User wants to give bad feedback to mindfulness exercises   |   |
| Steps                 | Wysa   | User  |
| 1                     |  | "Give me some mindfulness exercises for Asian elders" |
| 2                     | Wysa offers some personalized mindfulness exercises  |   |
| 3                     |  | User chooses one of the exercises                     |
| 4                     | Wysa guides user through the exercise  | User follows the exercises                            |
| 5                     |  | "This exercise was not helpful for me"                |
| 6                     | Wysa was able to recognize the conversation overall was not helpful but does not realize it was about the exercise |   |
| Expected output       | <ul style="list-style-type: none"> <li>Wysa should be able to guide user through exercises and</li> </ul>          |   |

|  |                          |
|--|--------------------------|
|  | incorporate the feedback |
|--|--------------------------|

Test 5:

|                       |  |   |
|-----------------------|--|---|
| Test case ID          | E5   |   |
| Test case description | User wants to do reframe thoughts (once)   |   |
| Steps                 | Wysa   | User  |
| 1                     |  | "I want to reframe thoughts to feel better as a Latino kid" |
| 2                     | Wysa guides user through reframe thoughts  | User follows it and reframe their thoughts                  |
| 3                     |  | "Give me the thoughts that I change"                        |
| Expected output       | <ul style="list-style-type: none"> <li>• Wysa should be able to guide user through the process and remember what had been discussed</li> </ul> |   |

Test 6:

|                       |   |   |
|-----------------------|---|---|
| Test case ID          | E6  |   |
| Test case description | User wants to do set schedule for reframe thoughts  |   |
| Steps                 | Wysa  | User  |
| 1                     |   | "As a Latino kid, I want to set schedule to reframe thoughts every Wednesday" |
| 2                     | Wysa guides user through the schedule set up  | User follows the setup  |
| Expected output       | <ul style="list-style-type: none"> <li>• Wysa should be able to guide user through the routine</li> </ul> |   |

Test 7:

User wants to give good feedback for reframe thoughts

|                       |  |  |
|-----------------------|--|--|
| Test case ID          | E7   |  |
| Test case description | User wants to give good feedback for reframe thoughts  |  |
| Steps                 | Wysa   | User   |
| 1                     |  | "I want to reframe thoughts to feel better as an Asian teen" |
| 2                     | Wysa guides user through reframe thoughts  | User follows it and reframe their thoughts                   |
| 3                     |  | "The exercise was helpful"                                   |
| Expected output       | <ul style="list-style-type: none"> <li>• Wysa should be able to guide user through reframing thoughts process, offer some personalization and incorporated the feedback</li> </ul> |  |

#### Test 8:

User wants to give bad feedback for reframe thoughts

|                       |  |  |
|-----------------------|--|--|
| Test case ID          | E8   |  |
| Test case description | User wants to do reframe thoughts (once)   |  |
| Steps                 | Wysa   | User   |
| 1                     |  | "I want to reframe thoughts to feel better as an Asian teen" |
| 2                     | Wysa guides user through reframe thoughts  | User follows it and reframe their thoughts                   |
| 3                     |  | "The exercise was unhelpful"                                 |
| Expected output       | <ul style="list-style-type: none"> <li>• Wysa should be able to guide user through reframing thoughts process and incorporated the feedback for later</li> </ul> |  |

#### Test 9:

User wants to do physical strength exercise (once)

|                       |   |  |
|-----------------------|---|--|
| Test case ID          | E9  |  |
| Test case description | User wants to do strength exercises (once)  |  |
| Steps                 | Wysa  | User   |
| 1                     |   | “Give me some strength exercises for Asian elders” |
| 2                     | Wysa offers some specialized strength exercises   |  |
| 3                     |   | User chooses one of the exercises                  |
| 4                     | Wysa guides user through the exercise   | User follows the exercises                         |
| 5                     |   | User asks what exercises been done                 |
| Expected output       | <ul style="list-style-type: none"> <li>Wysa should be able to guide user through exercises and remember what exercises been done</li> </ul> |  |

#### Test 10:

User wants to set routine for physical strength exercise

|                       |  |   |
|-----------------------|--|---|
| Test case ID          | E10  |   |
| Test case description | User wants to set exercise routines for physical strength exercises  |   |
| Steps                 | Wysa   | User  |
| 1                     |  | “I want to set some strength exercise routines as an Asian elder” |
| 2                     | Wysa guides user the exercise routine process  | User follows the exercise routine process                         |
| 3                     |  | User asks what exercises routines been set up                     |
| Expected output       | <ul style="list-style-type: none"> <li>Wysa should be able to guide user through exercise routines and remember the routines set up</li> </ul> |   |

Test 11:

User wants to give good feedback for physical strength exercise

|                       |  |  |
|-----------------------|--|--|
| Test case ID          | E11  |  |
| Test case description | User wants to give good feedback to physical strength exercises  |  |
| Steps                 | Wysa   | User   |
| 1                     |  | “Give me some strength exercises for Asian elders” |
| 2                     | Wysa offers some personalized physical exercises   |  |
| 3                     |  | User chooses one of the exercises                  |
| 4                     | Wysa guides user through the exercise  | User follows the exercises                         |
| 5                     |  | “This exercise was helpful for me”                 |
| Expected output       | <ul style="list-style-type: none"><li>• Wysa should be able to guide user through exercises and incorporate the feedback</li></ul> |  |

Test 12:

User wants to give bad feedback for physical strength exercise

|                       |   |  |
|-----------------------|---|--|
| Test case ID          | E12   |  |
| Test case description | User wants to give bad feedback to strength exercises |  |
| Steps                 | Wysa  | User   |
| 1                     |   | “Give me some strength exercises for Asian elders” |
| 2                     | Wysa offers some personalized strength exercises      |  |
| 3                     |   | User chooses one of the exercises                  |

|                 |  |  |
|-----------------|--|--|
| 4               | Wysa guides user through the exercise  | User follows the exercises             |
| 5               |  | "This exercise was not helpful for me" |
| Expected output | <ul style="list-style-type: none"> <li>Wysa should be able to guide user through exercises and incorporate the feedback</li> </ul> |  |

#### Test 13:

User wants to do durability exercise (once)

|                       |   |                                    |
|-----------------------|---|------------------------------------|
| Test case ID          | E13   |                                    |
| Test case description | User wants to do durability exercises (once)  |                                    |
| Steps                 | Wysa  | User                               |
| 1                     | "Give me some durability exercises for unemployed adult"  |                                    |
| 2                     | Wysa offers some physical exercises   |                                    |
| 3                     |   | User chooses one of the exercises  |
| 4                     | Wysa guides user through the exercise   | User follows the exercises         |
| 5                     |   | User asks what exercises been done |
| Expected output       | <ul style="list-style-type: none"> <li>Wysa should be able to guide user through exercises and remember what exercises been done</li> </ul> |                                    |

#### Test 14:

User wants to set routine for durability exercise

|                       |   |      |
|-----------------------|---|------|
| Test case ID          | E14   |      |
| Test case description | User wants to set exercise routines for mindfulness exercises |      |
| Steps                 | Wysa  | User |

|                 |  |   |
|-----------------|--|---|
| 1               |  | "I want to set routine to do mindfulness every Monday as an Latino adult" |
| 2               | Wysa guides user the exercise routine process  | User follows the exercise routine process                                 |
| 3               |  | User asks what exercises routines been set up                             |
| Expected output | <ul style="list-style-type: none"> <li>• Wysa should be able to guide user through exercise routines and remember the routines set up</li> </ul> |   |

Test 15:

User wants to give good feedback for durability exercise

|                       |  |                                    |
|-----------------------|--|------------------------------------|
| Test case ID          | E15  |                                    |
| Test case description | User wants to give good feedback to mindfulness exercises  |                                    |
| Steps                 | Wysa   | User                               |
| 1                     | "Give me some mindfulness exercises for Asian elders"  |                                    |
| 2                     | Wysa offers some personalized mindfulness exercises  |                                    |
| 3                     |  | User chooses one of the exercises  |
| 4                     | Wysa guides user through the exercise  | User follows the exercises         |
| 5                     |  | "This exercise was helpful for me" |
| Expected output       | <ul style="list-style-type: none"> <li>• Wysa should be able to guide user through exercises and incorporate the feedback</li> </ul> |                                    |

Test 16:

User wants to give bad feedback for durability exercise

|              |     |
|--------------|-----|
| Test case ID | E16 |
|--------------|-----|

|                       |  |  |
|-----------------------|--|--|
| Test case description | User wants to give bad feedback to durability exercises  |  |
| Steps                 | Wysa   | User   |
| 1                     |  | “Give me some durability exercises for Asian elders” |
| 2                     | Wysa offers some personalized durability exercises   |  |
| 3                     |  | User chooses one of the exercises                    |
| 4                     | Wysa guides user through the exercise  | User follows the exercises                           |
| 5                     |  | “This exercise was not helpful for me”               |
| Expected output       | <ul style="list-style-type: none"> <li>• Wysa should be able to offer personalized mindfulness exercises and incorporate the feedback</li> </ul> |  |

Test 17:  
Ask different languages

|                       |   |   |
|-----------------------|---|---|
| Test case ID          | E17   |   |
| Test case description | User wants to give feedback to physical exercises   |   |
| Steps                 | Wysa  | User  |
| 1                     |   | “Tôi muốn tập thể dục” (I want to exercise but in Vietnamese) |
| 2                     | Wysa should either:<br>1. Respond in the appropriate language<br>2. Tell the user to respond in English as it currently does not support this language or |   |

|                 |  |  |
|-----------------|--|--|
|                 | other language than English  |  |
| Expected output | Wysa should be able to detect different language and asks to clarify |  |

Test 18:  
Ask in improper conversation format

|                       |  |  |
|-----------------------|--|--|
| Test case ID          | E18  |  |
| Test case description | User wants to change exercises to physical exercises   |  |
| Steps                 | Wysa   | User   |
| 1                     |  | User asks to do mindfulness exercises                                |
| 2                     | Wysa offers some mindfulness exercises   |  |
| 3                     |  | User chooses one of the exercises                                    |
| 4                     | Wysa guides user through the exercise  | User follows the exercises   |
| 5                     |  | "I want to do physical exercises instead"                            |
| 6                     | Wysa offers some physical exercises  |  |
| 7                     |  | User chooses one of the physical exercises                           |
| 8                     | Wysa guides user through the exercises   | User keeps asking to change exercises instead of doing the exercises |
| 9                     | Wysa should be able to detect the improper conversation format and detect based on the outline |  |

|                 |  |
|-----------------|--|
| Expected output | Wysa should be able to detect the improper conversation format and detect based on the outline |
|-----------------|--|

### Augmented test case

For AI testing augmented test cases, we will use test case 1 as the base case

| ID | Augment Profile       | Augment Result  | Wysa able to recognize request for mindfulness exercise | Wysa offers exercise personalization |
|----|-----------------------|---|---|--------------------------------------|
| 1  | random word swap      | Give some me mindfulness exercises for Asian elders             | Expected to   | Expected to                          |
| 2  | random word delete    | me some exercises for Asian elders                              | Expected to   | Expected to                          |
| 3  | random word insert    | Give heedfulness me some mindfulness exercises for Asian elders | Expected to   | Expected to                          |
| 4  | synonym augmentation  | Afford me some heedfulness exercises for Asian elders           | Expected to   | Expected to                          |
| 5  | ocr augmentation      | Give me some mindfolne88 exercises for Asian eldeks             | Expected to   | Expected to                          |
| 6  | keyboard augmentation | Give me some mindfulness exe3dises for Asian #lders             | Expected to   | Expected to                          |
| 7  | random char insert    | wGive me some mindfulness exercises for                         | Expected to   | Expected to                          |

|   |                    |   |             |             |
|---|--------------------|---|-------------|-------------|
|   |                    | A2sian elders                                       |             |             |
| 8 | random char swap   | Give me some mindfulness exercises for Asian elders | Expected to | Expected to |
| 9 | random char delete | Give me some mindfulness exercises for Asian elders | Expected to | Expected to |

## Invalid Topics Function

Decision Table

| <b>id</b> | <b>Not mental health related.Topic</b> | <b>Not mental health related.Type</b> |
|-----------|--|---------------------------------------|
| 1         | Cooking                                | Question                              |
| 2         | Stock Market                           | Statement                             |
| 3         | Cooking                                | Statement                             |
| 4         | Sports                                 | Question                              |
| 5         | Sports                                 | Statement                             |
| 6         | Weather                                | Question                              |
| 7         | Weather                                | Statement                             |
| 8         | Math                                   | Question                              |
| 9         | Math                                   | Statement                             |
| 10        | Stock Market                           | Question                              |

|   |                           |              |        |         |      |          |
|---|---------------------------|--------------|--------|---------|------|----------|
|   | Not mental health related |              |        |         |      |          |
|   | Topic                     |              |        |         |      | Type     |
|   | Cooking                   | Stock Market | Sports | Weather | Math | Question |
| 1 | X                         |              |        |         |      | X        |

|    |   |   |   |   |   |  |   |   |
|----|---|---|---|---|---|--|---|---|
| 2  | X |   |   |   |   |  |   | X |
| 3  |   | X |   |   |   |  | X |   |
| 4  |   | X |   |   |   |  |   | X |
| 5  |   |   | X |   |   |  | X |   |
| 6  |   |   | X |   |   |  |   | X |
| 7  |   |   |   | X |   |  | X |   |
| 8  |   |   |   | X |   |  |   | X |
| 9  |   |   |   |   | X |  | X |   |
| 10 |   |   |   |   | X |  |   | X |

Test case 1:

|                 |   |   |
|-----------------|---|---|
| Test case ID    | I1  |   |
| Input           | Topic: Cooking  | Type: Question                              |
| Steps           | Wysa  | User  |
| 1               | Ask how the user is doing   |   |
| 2               |   | How do you know if a burger is done or not? |
| 3               | Should be able to detect that the conversation is not about the user's mental state |   |
| Expected output | Should be able to detect that the conversation is not about the user's mental state |   |

Test case 2:

|              |                |                 |
|--------------|----------------|-----------------|
| Test case ID | I2             |                 |
| Input        | Topic: Cooking | Type: Statement |

| Steps           | Wysa  | User   |
|-----------------|---|--|
| 1               | Ask how the user is doing   |  |
| 2               |   | Chicken is safe to eat at 165 degrees Fahrenheit |
| 3               | Should be able to detect that the conversation is not about the user's mental state |  |
| Expected output | Should be able to detect that the conversation is not about the user's mental state |  |

Test case 3:

| Test case ID    | I3  |  |
|-----------------|---|--|
| Input           | Topic: Stock Market   | Type: Question                                 |
| Steps           | Wysa  | User   |
| 1               | Ask how the user is doing   |  |
| 2               |   | What is Downey's stock looking like right now? |
| 3               | Should be able to detect that the conversation is not about the user's mental state |  |
| Expected output | Should be able to detect that the conversation is not about the user's mental state |  |

Test case 4:

| Test case ID | I4                        |                 |
|--------------|---------------------------|-----------------|
| Input        | Topic: Stock Market       | Type: Statement |
| Steps        | Wysa                      | User            |
| 1            | Ask how the user is doing |                 |

|                 |   |   |
|-----------------|---|---|
| 2               |   | The general rule of thumb in the stock market is to buy low and sell high |
| 3               | Should be able to detect that the conversation is not about the user's mental state |   |
| Expected output | Should be able to detect that the conversation is not about the user's mental state |   |

Test case 5:

|                 |   |   |
|-----------------|---|---|
| Test case ID    | I5  |   |
| Input           | Topic: Sports   | Type: Question                            |
| Steps           | Wysa  | User                                      |
| 1               | Ask how the user is doing   |   |
| 2               |   | Why were the Washington Redskins renamed? |
| 3               | Should be able to detect that the conversation is not about the user's mental state |   |
| Expected output | Should be able to detect that the conversation is not about the user's mental state |   |

Test case 6:

|              |                           |  |
|--------------|---------------------------|--|
| Test case ID | I6                        |  |
| Input        | Topic: Sports             | Type: Statement  |
| Steps        | Wysa                      | User   |
| 1            | Ask how the user is doing |  |
| 2            |                           | American football is not the same as football in other |

|                 |   |           |
|-----------------|---|-----------|
|                 |   | countries |
| 3               | Should be able to detect that the conversation is not about the user's mental state |           |
| Expected output | Should be able to detect that the conversation is not about the user's mental state |           |

Test case 7:

|                 |   |                         |
|-----------------|---|-------------------------|
| Test case ID    | I7  |                         |
| Input           | Topic: Weather  | Type: Question          |
| Steps           | Wysa  | User                    |
| 1               | Ask how the user is doing   |                         |
| 2               |   | Will it rain next week? |
| 3               | Should be able to detect that the conversation is not about the user's mental state |                         |
| Expected output | Should be able to detect that the conversation is not about the user's mental state |                         |

Test case 8:

|              |   |  |
|--------------|---|--|
| Test case ID | I8  |  |
| Input        | Topic: Weather  | Type: Statement  |
| Steps        | Wysa  | User   |
| 1            | Ask how the user is doing                                       |  |
| 2            |   | California experienced turbulent weather changes in 2023 |
| 3            | Should be able to detect that the conversation is not about the |  |

|                 |   |  |
|-----------------|---|--|
|                 | user's mental state   |  |
| Expected output | Should be able to detect that the conversation is not about the user's mental state |  |

Test case 9:

|                 |   |                              |
|-----------------|---|------------------------------|
| Test case ID    | I9  |                              |
| Input           | Topic: Math   | Type: Question               |
| Steps           | Wysa  | User                         |
| 1               | Ask how the user is doing   |                              |
| 2               |   | What is the derivative of 8? |
| 3               | Should be able to detect that the conversation is not about the user's mental state |                              |
| Expected output | Should be able to detect that the conversation is not about the user's mental state |                              |

Test case 10:

|                 |   |                      |
|-----------------|---|----------------------|
| Test case ID    | I10   |                      |
| Input           | Topic: Math   | Type: Statement      |
| Steps           | Wysa  | User                 |
| 1               | Ask how the user is doing   |                      |
| 2               |   | 4 is the square of 2 |
| 3               | Should be able to detect that the conversation is not about the user's mental state |                      |
| Expected output | Should be able to detect that the conversation is not about the user's mental state |                      |

## Section 3.2 Test case reports

### Mental Illness Function

Test case 1: Recommends an exercise/technique

| Test Case ID          | M1  |               |                               |                       |  |  |
|-----------------------|---|---------------|-------------------------------|-----------------------|--|--|
| Test Case Description | User have anxiety and depression that is caused by negative thoughts and traumatic events.  |               |                               |                       |  |  |
| Input                 | Causes.Negative   | Causes.Trauma | Mental Issues.Disorder        | Mental Issues.Illness |  |  |
|                       | Thoughts  | Events        | Anxiety                       | Depression            |  |  |
| Steps                 | Wysa  |               | User                          |                       |  |  |
| 1                     | Asks User how they are feeling  |               |                               |                       |  |  |
| 2                     |   |               | Says, "I'm feeling depressed" |                       |  |  |
| 3                     | Asks User why they are depressed  |               |                               |                       |  |  |
| 4                     |   |               | Says, "I have anxiety"        |                       |  |  |
| 5                     | Recommends the "Reframe Thoughts" exercise  |               |                               |                       |  |  |
| 6                     |   |               | Does exercise                 |                       |  |  |
| Actual Output         | Instead of asking what thoughts the User has, Wysa recommends the "Reframe Thoughts" exercise that will reduce the intensity of emotions that the User's thoughts can affect. |               |                               |                       |  |  |
| Invalid/Valid         | Valid   |               |                               |                       |  |  |

Test case 2: Change in User's conversation format

| Test Case ID          | M2  |               |                        |                       |
|-----------------------|---|---------------|------------------------|-----------------------|
| Test Case Description | User has anxiety and depression that is caused by negative feelings and traumatic events. |               |                        |                       |
| Input                 | Causes.Negative   | Causes.Trauma | Mental Issues.Disorder | Mental Issues.Illness |

|               | Feelings   | Events | Anxiety  | Depression |
|---------------|--|--------|--|------------|
| Steps         | Wysa   |        | User   |            |
| 1             | Asks how the User is feeling   |        |  |            |
| 2             |  |        | Says, "What are different exercises that will help with my anxiety?" |            |
| 3             | Asks, "Can I help you with your feelings of anxiety?"  |        |  |            |
| 4             | Prompts User to choose, "I'd like that", "Let's keep talking", "You misunderstood me"  |        |  |            |
| 5             |  |        | Presses, "You misunderstood me"                                      |            |
| 6             | Asks, "Shall we go back to what we were talking about?   |        |  |            |
| Actual Output | Wysa was able to understand that the User had anxiety. However, Wysa didn't provide different exercises that will help the User. |        |  |            |
| Invalid/Valid | Invalid  |        |  |            |

### Test case 3: Short Term Memory

|                       |  |               |                             |                       |
|-----------------------|--|---------------|-----------------------------|-----------------------|
| Test Case ID          | M3   |               |                             |                       |
| Test Case Description | User is a student that is feeling stressed because of too much homework and creates a positive thought |               |                             |                       |
| Input                 | Causes.Negative  | Causes.Trauma | Mental Issues.Disorder      | Mental Issues.Illness |
|                       | Thoughts   | Events        | Anxiety                     | Stress                |
| Steps                 | Wysa   |               | User                        |                       |
| 1                     | Asks User how they are feeling   |               |                             |                       |
| 2                     |  |               | Says "I'm feeling stressed" |                       |

|               |   |                                      |
|---------------|---|--------------------------------------|
| 3             | Asks if User is also anxious  |                                      |
| 4             |   | Clicks "Yes"                         |
| 5             | Asks if it is affecting the User physically too   |                                      |
| 6             | Prompts User to choose, "Mostly mind", "Mostly physical", or "A bit of both"  |                                      |
| 7             |   | User chooses "Mostly mind"           |
| 8             | Asks User what thought comes to mind that causes their stress   |                                      |
| 9             |   | User says "I have a lot of homework" |
| 10            | Asks the user what change they would hope for   |                                      |
| 11            |   | User says, "I want to be happy"      |
| 12            | Restates the user's thought, "I have a lot of homework" and asks what positives about the situation/themselves have been ignored                        |                                      |
| 13            |   | User says, "I am smart"              |
| 14            | Restates User's positive thought and congratulates the User for creating a positive thought   |                                      |
| Actual Output | Wysa was able to remember the User's mental health and thought input. Lead the User to create a positive thought that will help boost their confidence. |                                      |
| Invalid/Valid | Valid   |                                      |

#### Test case 4: Long Term Memory

|              |    |
|--------------|----|
| Test Case ID | M4 |
|--------------|----|

|                       |  |               |  |                       |
|-----------------------|--|---------------|--|-----------------------|
| Test Case Description | User is a student that is feeling stressed because of a midterm and starts a new conversation after creating a positive thought. |               |  |                       |
| Input                 | Causes.Negative  | Causes.Trauma | Mental Issues.Disorder                   | Mental Issues.Illness |
|                       | Feelings   | Events        | Anxiety                                  | Stress                |
| Steps                 | Wysa   |               | User                                     |                       |
| 1                     | Asks User how they are feeling   |               |  |                       |
| 2                     |  |               | Says "I'm feeling stressed"              |                       |
| 3                     | Asks if User is also anxious   |               |  |                       |
| 4                     |  |               | Clicks "Yes"                             |                       |
| 5                     | Asks if it is affecting the User physically too  |               |  |                       |
| 6                     | Prompts User to choose, "Mostly mind", "Mostly physical", or "A bit of both"   |               |  |                       |
| 7                     |  |               | User chooses "Mostly mind"               |                       |
| 8                     | Asks User what thought comes to mind that causes their stress  |               |  |                       |
| 9                     |  |               | User says "I have a midterm today"       |                       |
| 10                    | Asks the user what change they would hope for  |               |  |                       |
| 11                    |  |               | User says what they would like to change |                       |
| 12                    | Restates the user's thought, "I have a lot of homework" and asks what positives about the situation/themselves have been ignored |               |  |                       |
| 13                    |  |               | User says, "I want to be happy"          |                       |

|               |  |                               |
|---------------|--|-------------------------------|
| 14            | Restates the user's thought, "I have a lot of homework" and asks what positives about the situation/themselves have been ignored |                               |
| 15            |  | User says, "I am smart"       |
| 16            | Restates User's positive thought and congratulates the User for creating a positive thought                                      |                               |
| 17            | Prompts the User to "Start new Chat"   |                               |
| 18            |  | User presses "Start new Chat" |
| 19            | Wysa states that User has mentioned that they're "feeling stressed"  |                               |
| Actual Output | Wysa was able to remember that the User was feeling stressed after the User created a new chat.                                  |                               |
| Invalid/Valid | Valid  |                               |

#### Test case 5: Recommends Sleep Exercises

|                       |   |               |                               |                       |
|-----------------------|---|---------------|-------------------------------|-----------------------|
| Test Case ID          | M5  |               |                               |                       |
| Test Case Description | User needs help to sleep.                       |               |                               |                       |
| Input                 | Causes.Negative                                 | Causes.Trauma | Mental Issues.Disorder        | Mental Issues.Illness |
|                       | Thoughts  | Events        | Sleep                         | Depression            |
| Steps                 | Wysa  |               | User                          |                       |
| 1                     |   |               | Says, "I am lacking in sleep" |                       |
| 2                     | Asks User if they would like to help them sleep |               |                               |                       |
| 3                     |   |               | Clicks, "Yes"                 |                       |

|               |  |  |
|---------------|--|--|
| 4             | Asks if they are also feeling tired  |  |
| 5             |  | Can choose, "Yes" or "Not really"                  |
| 6             | Asks User if they would like to try a meditation to help them sleep                        |  |
| 7             |  | Clicks, "No"                                       |
| 8             | Provides all sleep meditations and relaxation exercises                                    |  |
| 9             |  | User chooses a sleep recommendation and follows it |
| Actual Output | Wysa was able to provide sleep meditation and relaxation exercises to help the User sleep. |  |
| Invalid/Valid | Valid  |  |

#### Test case 6: Improper Conversation Format

|                       |  |               |   |                       |
|-----------------------|--|---------------|---|-----------------------|
| Test Case ID          | M6   |               |   |                       |
| Test Case Description | User needs help staying awake and wants to know why they are depressed and having trouble sleeping                                 |               |   |                       |
| Input                 | Causes.Negative  | Causes.Trauma | Mental Issues.Disorder                    | Mental Issues.Illness |
|                       | Feelings   | Events        | Sleep                                     | Depression            |
| Steps                 | Wysa   |               | User                                      |                       |
| 1                     |  |               | Asks, "Why am I having trouble sleeping?" |                       |
| 2                     | Advises User to check the self-care section for sleep stories  |               |   |                       |
| 3                     |  |               | Asks, "Can you help me stay up?"          |                       |
| 4                     | Provides a slider with quick features that contain, "Privacy", "Relax and Calm", "SOS", "Emotive Reflections" and advises the User |               |   |                       |

|                 |   |                             |
|-----------------|---|-----------------------------|
|                 | to check out the self-care kit  |                             |
| 5               |   | Asks, "Why am I depressed?" |
| 6               | Asks what happened that made the user feel depressed  |                             |
| Expected Output | Wysa should be able to provide the right recommendations to the user and ask questions that will help determine the User's mental condition.  |                             |
| Actual Output   | Wysa was able to provide information about the self-care section that contains sleep stories. Wysa provides quick features that were irrelevant to the User's question, "Can you help me stay up?" Wysa was able to recognize the User's question, "Why am I depressed?" and asks why they are depressed. |                             |
| Invalid/Valid   | Invalid (because Wysa did not answer one question right)  |                             |

#### Test case 7: Different Language

|                       |   |               |                        |                       |
|-----------------------|---|---------------|------------------------|-----------------------|
| Test Case ID          | M7  |               |                        |                       |
| Test Case Description | User tells Wysa their mental illness in a different language.                         |               |                        |                       |
| Input                 | Causes.Negative   | Causes.Trauma | Mental Issues.Disorder | Mental Issues.Illness |
|                       | Thoughts  | Events        | Sleep                  | Stress                |
| Steps                 | Wysa  |               | User                   |                       |
| 1                     | Asks how the User is feeling  |               |                        |                       |
| 2                     |   |               | Says, "ストレスがたまります"     |                       |
| 3                     | Says, "That looks like a bunch of letters that don't make a lot of sense together. 😕" |               |                        |                       |
| 4                     | Asks, "Are you testing my AI capabilities   |               |                        |                       |
| 5                     | Prompts for User to choose, "No, I'm not", "Yes", "It's my language"                  |               |                        |                       |

|               |  |                                  |
|---------------|--|----------------------------------|
| 6             |  | User presses, “It’s my language” |
| 7             | Says, “I can only respond in English right now. I hope to learn more languages soon.”          |                                  |
| Actual Output | Wysa wasn’t able to understand the User’s Japanese language. Wysa can only respond in English. |                                  |
| Invalid/Valid | Invalid  |                                  |

Test case 8: Text Data Augmentation

Test case 8.0: Normal Text

|                       |   |               |  |                       |  |  |
|-----------------------|---|---------------|--|-----------------------|--|--|
| Test Case ID          | M8.0  |               |  |                       |  |  |
| Test Case Description | User states that they are stressed out.   |               |  |                       |  |  |
| Input                 | Causes.Negative   | Causes.Trauma | Mental Issues.Disorder                     | Mental Issues.Illness |  |  |
|                       | Feelings  | Events        | Sleep                                      | Stress                |  |  |
| Steps                 | Wysa  |               | User                                       |                       |  |  |
| 1                     | Asks how is the User’s day  |               |  |                       |  |  |
| 2                     |   |               | User responds “I am feeling stressed out.” |                       |  |  |
| 3                     | States that this feeling brings a lot of unpleasantness and advises User to take this as a positive push. |               |  |                       |  |  |
| 4                     | Asks User why they are feeling stressed   |               |  |                       |  |  |
| Actual Output         | Wysa was able to recognize that the User is stressed and ask why they are feeling that way.               |               |  |                       |  |  |
| Invalid/Valid         | Valid   |               |  |                       |  |  |

Test case 8.1: Random Word Swap

|                       |  |               |  |                       |  |  |
|-----------------------|--|---------------|--|-----------------------|--|--|
| Test Case ID          | M8.2   |               |  |                       |  |  |
| Test Case Description | Original Text Input: I am feeling stressed out<br>Text data augmentation type: random word swap<br>Step 2 is being augmented |               |  |                       |  |  |
| Input                 | Causes.Negative  | Causes.Trauma | Mental Issues.Disorder                     | Mental Issues.Illness |  |  |
|                       | Feelings   | Events        | Sleep                                      | Stress                |  |  |
| Steps                 | Wysa   |               | User                                       |                       |  |  |
| 1                     | Asks how is the User's day   |               |  |                       |  |  |
| 2                     |  |               | User responds "I out. feeling stressed am" |                       |  |  |
| 3                     | Says, "Do help me understand you better. Tell me what isn't working for you."  |               |  |                       |  |  |
| Actual Output         | Wysa wasn't able to understand the User but tells the User to tell Wysa what isn't working for them.                         |               |  |                       |  |  |
| Invalid/Valid         | Invalid  |               |  |                       |  |  |

Test case 8.2: Random Word Delete

|                       |  |               |                                 |                       |
|-----------------------|--|---------------|---------------------------------|-----------------------|
| Test Case ID          | M8.2   |               |                                 |                       |
| Test Case Description | Original Text Input: I am feeling stressed out<br>Text data augmentation type: random word delete<br>Step 2 is being augmented |               |                                 |                       |
| Input                 | Causes.Negative  | Causes.Trauma | Mental Issues.Disorder          | Mental Issues.Illness |
|                       | Feelings   | Events        | Sleep                           | Stress                |
| Steps                 | Wysa   |               | User                            |                       |
| 1                     | Asks how is the User's day   |               |                                 |                       |
| 2                     |  |               | User responds "am feeling out." |                       |

|               |  |  |
|---------------|--|--|
| 3             | Says, “Oh...<br>Tell me more. You may feel better talking about it.”                       |  |
| Actual Output | Wysa wasn't able to understand the User but tells the User to tell Wysa to tell them more. |  |
| Invalid/Valid | Invalid  |  |

#### Test case 8.3: Random Word Insert

|                       |  |               |  |                       |  |  |
|-----------------------|--|---------------|--|-----------------------|--|--|
| Test Case ID          | M8.3   |               |  |                       |  |  |
| Test Case Description | Original Text Input: I am feeling stressed out<br>Text data augmentation type: random word insert<br>Step 2 is being augmented |               |  |                       |  |  |
| Input                 | Causes.Negative  | Causes.Trauma | Mental Issues.Disorder                                       | Mental Issues.Illness |  |  |
|                       | Feelings   | Events        | Sleep  | Stress                |  |  |
| Steps                 | Wysa   |               | User   |                       |  |  |
| 1                     | Asks how is the User's day   |               |  |                       |  |  |
| 2                     |  |               | User responds “I am feeling stressed intuitive feeling out.” |                       |  |  |
| 3                     | Says, “Oh...<br>Tell me more. You may feel better talking about it.”   |               |  |                       |  |  |
| Actual Output         | Wysa wasn't able to understand the User but tells the User to tell Wysa to tell them more.                                     |               |  |                       |  |  |
| Invalid/Valid         | Invalid  |               |  |                       |  |  |

#### Test case 8.4: Synonym Augmentation

|                       |  |               |        |        |
|-----------------------|--|---------------|--------|--------|
| Test Case ID          | M8.4   |               |        |        |
| Test Case Description | Original Text Input: I am feeling stressed out<br>Text data augmentation type: synonym augmentation<br>Step 2 is being augmented |               |        |        |
| Input                 | Causes.Negative  | Causes.Trauma | Mental | Mental |

|               |   |        |  |                |  |  |
|---------------|---|--------|--|----------------|--|--|
|               |   | a      | Issues.Disorder                            | Issues.Illness |  |  |
|               | Feelings  | Events | Sleep                                      | Stress         |  |  |
| Steps         | Wysa  |        | User                                       |                |  |  |
| 1             | Asks how is the User's day  |        |  |                |  |  |
| 2             |   |        | User responds "I am feeling stressed out." |                |  |  |
| 3             | Says, "That's an uncomfortable feeling, I know. Observe how it feels for a bit, just allowing it to exist." |        |  |                |  |  |
| 4             | Asks User what is making them feel this way   |        |  |                |  |  |
| Actual Output | Wysa was able to understand the User's feelings and asks why they feel that way.                            |        |  |                |  |  |
| Invalid/Valid | Valid   |        |  |                |  |  |

#### Test case 8.5: OCR Augmentation

|                       |  |               |  |                       |
|-----------------------|--|---------------|--|-----------------------|
| Test Case ID          | M8.5   |               |  |                       |
| Test Case Description | Original Text Input: I am feeling stressed out<br>Text data augmentation type: ocr augmentation<br>Step 2 is being augmented |               |  |                       |
| Input                 | Causes.Negative  | Causes.Trauma | Mental Issues.Disorder                     | Mental Issues.Illness |
|                       | Feelings   | Events        | Sleep                                      | Stress                |
| Steps                 | Wysa   |               | User                                       |                       |
| 1                     | Asks how is the User's day   |               |  |                       |
| 2                     |  |               | User responds "I am feeling stressed out." |                       |
| 3                     | Says, "That's an uncomfortable feeling, I know. Observe how it   |               |  |                       |

|               |   |  |
|---------------|---|--|
|               | feels for a bit, just allowing it to exist.”                                      |  |
| 4             | Asks User what is making them feel this way                                       |  |
| Actual Output | Wysa was able to understand the User’s feelings and asked why they felt that way. |  |
| Invalid/Valid | Valid   |  |

#### Test case 8.6: Keyboard Augmentation

|                       |   |               |  |                       |  |  |
|-----------------------|---|---------------|--|-----------------------|--|--|
| Test Case ID          | M8.6  |               |  |                       |  |  |
| Test Case Description | Original Text Input: I am feeling stressed out<br>Text data augmentation type: keyboard augmentation<br>Step 2 is being augmented |               |  |                       |  |  |
| Input                 | Causes.Negative   | Causes.Trauma | Mental Issues.Disorder                     | Mental Issues.Illness |  |  |
|                       | Feelings  | Events        | Sleep                                      | Stress                |  |  |
| Steps                 | Wysa  |               | User                                       |                       |  |  |
| 1                     | Asks how is the User’s day  |               |  |                       |  |  |
| 2                     |   |               | User responds “I am feeling stressed out.” |                       |  |  |
| 3                     | Says, “Oh! Tell me more about that”   |               |  |                       |  |  |
| Actual Output         | Wysa wasn’t able to understand the User’s feelings but told the user to tell them more about what they said.                      |               |  |                       |  |  |
| Invalid/Valid         | Invalid   |               |  |                       |  |  |

#### Test case 8.7: Random Char Insert

|                       |  |
|-----------------------|--|
| Test Case ID          | M8.7   |
| Test Case Description | Original Text Input: I am feeling stressed out<br>Text data augmentation type: random char insert<br>Step 2 is being augmented |

|               |  |               |  |                       |  |  |
|---------------|--|---------------|--|-----------------------|--|--|
| Input         | Causes.Negative  | Causes.Trauma | Mental Issues.Disorder                     | Mental Issues.Illness |  |  |
|               | Feelings   | Events        | Sleep                                      | Stress                |  |  |
| Steps         | Wysa   |               | User                                       |                       |  |  |
| 1             | Asks how is the User's day   |               |  |                       |  |  |
| 2             |  |               | User responds "I am feeling stressed out." |                       |  |  |
| 3             | Says, "Oh, I'm here for you. Tell me more about what happened."  |               |  |                       |  |  |
| Actual Output | Wysa wasn't able to understand the User's feelings but told the user to tell them more about what they said. |               |  |                       |  |  |
| Invalid/Valid | Invalid  |               |  |                       |  |  |

Test case 8.8: Random Char Swap

|                       |  |               |  |                       |
|-----------------------|--|---------------|--|-----------------------|
| Test Case ID          | M8.8   |               |  |                       |
| Test Case Description | Original Text Input: I am feeling stressed out<br>Text data augmentation type: random char swap<br>Step 2 is being augmented               |               |  |                       |
| Input                 | Causes.Negative  | Causes.Trauma | Mental Issues.Disorder                     | Mental Issues.Illness |
|                       | Feelings   | Events        | Sleep                                      | Stress                |
| Steps                 | Wysa   |               | User                                       |                       |
| 1                     | Asks how is the User's day   |               |  |                       |
| 2                     |  |               | User responds "I am feleign stressed out." |                       |
| 3                     | Says, "I understand how unpleasant emotions can disrupt our day to day life. How do you think this is impacting the way you see yourself?" |               |  |                       |

|               |  |
|---------------|--|
| Actual Output | Wysa was able to understand what the User was feeling and asks a valid question. |
| Invalid/Valid | Valid  |

Test case 8.9: Random Char Delete

|                       |  |               |  |                       |  |  |
|-----------------------|--|---------------|--|-----------------------|--|--|
| Test Case ID          | M8.9   |               |  |                       |  |  |
| Test Case Description | Original Text Input: I am feeling stressed out<br>Text data augmentation type: random char delete<br>Step 2 is being augmented             |               |  |                       |  |  |
| Input                 | Causes.Negative  | Causes.Trauma | Mental Issues.Disorder                     | Mental Issues.Illness |  |  |
|                       | Feelings   | Events        | Sleep                                      | Stress                |  |  |
| Steps                 | Wysa   |               | User                                       |                       |  |  |
| 1                     | Asks how is the User's day   |               |  |                       |  |  |
| 2                     |  |               | User responds "I am feeling stressed out." |                       |  |  |
| 3                     | Says, "I understand how unpleasant emotions can disrupt our day to day life. How do you think this is impacting the way you see yourself?" |               |  |                       |  |  |
| Actual Output         | Wysa was able to understand what the User was feeling and asks a valid question.   |               |  |                       |  |  |
| Invalid/Valid         | Valid  |               |  |                       |  |  |

Exercise function

Test case 1:

|              |    |
|--------------|----|
| Test case ID | E1 |
|--------------|----|

|                       |   |   |
|-----------------------|---|---|
| Test case description | User wants to do mindfulness exercises  |   |
| Steps                 | Wysa  | User  |
| 1                     |   | “Give me some mindfulness exercises for Asian elders” |
| 2                     | Wysa offers mindfulness exercises but without any specialization  |   |
| 3                     |   | User chooses one of the exercises                     |
| 4                     | Wysa guides user through the exercise   | User follows the exercises                            |
| 5                     |   | “What exercises did I do”                             |
| 6                     | Wysa was not able to understand user question and recommend new exercises instead   |   |
| Actual output         | <ul style="list-style-type: none"> <li>• Wysa was able to recognize the request for mindfulness exercises but no specialization</li> <li>• Wysa was not able to answer memory question</li> </ul> |   |

## Test case 2

|                       |  |   |
|-----------------------|--|---|
| Test case ID          | E2   |   |
| Test case description | User wants to set exercise routines for mindfulness exercises  |   |
| Steps                 | Wysa   | User  |
| 1                     |  | “I want to set some exercise routines as an Latino adult” |
| 2                     | Wysa wants to offer some physical activity   |   |
| Actual output         | <ul style="list-style-type: none"> <li>• Wysa was not able to recognize the request for routine exercises</li> </ul> |   |

### Test case 3

|                       |   |   |
|-----------------------|---|---|
| Test case ID          | E3  |   |
| Test case description | User wants to give good feedback to mindfulness exercises   |   |
| Steps                 | Wysa  | User  |
| 1                     |   | "Give me some mindfulness exercises for Asian elders" |
| 2                     | Wysa offers some personalized mindfulness exercises   |   |
| 3                     |   | User chooses one of the exercises                     |
| 4                     | Wysa guides user through the exercise   | User follows the exercises                            |
| 5                     |   | "This exercise was helpful for me"                    |
| 6                     | Wysa was able to recognize the conversation overall was not helpful but does not realize it was about the exercise                                      |   |
| Actual output         | <ul style="list-style-type: none"> <li>Wysa offers thumbs up and down after completing an exercise but does not realize the input after that</li> </ul> |   |

### Test case 4

|                       |  |   |
|-----------------------|--|---|
| Test case ID          | E4   |   |
| Test case description | User wants to give bad feedback to mindfulness exercises |   |
| Steps                 | Wysa   | User  |
| 1                     |  | "Give me some mindfulness exercises for Asian elders" |

|               |   |  |
|---------------|---|--|
| 2             | Wysa offers some personalized mindfulness exercises   |  |
| 3             |   | User chooses one of the exercises      |
| 4             | Wysa guides user through the exercise   | User follows the exercises             |
| 5             |   | "This exercise was not helpful for me" |
| 6             | Wysa was able to recognize the conversation overall was not helpful but does not realize it was about the exercise                                      |  |
| Actual output | <ul style="list-style-type: none"> <li>Wysa offers thumbs up and down after completing an exercise but does not realize the input after that</li> </ul> |  |

#### Test case 5

|                       |  |   |
|-----------------------|--|---|
| Test case ID          | E5   |   |
| Test case description | User wants to do reframe thoughts (once)   |   |
| Steps                 | Wysa   | User  |
| 1                     |  | "I want to reframe thoughts to feel better as a Latino kid" |
| 2                     | "Unhelpful thoughts can sap our energy and affect the way we see things."<br>"Would you like to find a fresh perspective with me?"               | User follows it and reframe their thoughts                  |
| 3                     |  | "Give me the thoughts that I change"                        |
| Actual output         | <ul style="list-style-type: none"> <li>Wysa guided user through the process</li> <li>Did remember what had been discussed without the</li> </ul> |   |

|  |   |
|--|---|
|  | <p>questions</p> <ul style="list-style-type: none"> <li>Offer no personalization for Latino or kid</li> </ul> |
|--|---|

#### Test case 6

|                       |  |   |
|-----------------------|--|---|
| Test case ID          | E6   |   |
| Test case description | User wants to do set schedule for reframe thoughts   |   |
| Steps                 | Wysa   | User  |
| 1                     |  | “As a Latino kid, I want to set schedule to reframe thoughts every Wednesday” |
| 2                     | “Unhelpful thoughts can sap our energy and affect the way we see things.”<br>“Would you like to find a fresh perspective with me?”                   |   |
| Actual output         | <ul style="list-style-type: none"> <li>Wysa was not able to recognize the request for scheduling and mistaken for just reframing thoughts</li> </ul> |   |

#### Test case 7

|                       |   |  |
|-----------------------|---|--|
| Test case ID          | E7  |  |
| Test case description | User wants to give good feedback for reframe thoughts |  |
| Steps                 | Wysa  | User   |
| 1                     |   | “I want to reframe thoughts to feel better as an Asian teen” |
| 2                     | Wysa guides user through reframe thoughts             | User follows it and reframe their thoughts                   |

|               |  |   |
|---------------|--|---|
| 3             |  | “The exercise was helpful”  |
| Actual output |  | <ul style="list-style-type: none"> <li>• Wysa was able to guide user through the reframing thoughts process</li> <li>• Wysa did not offer personalization</li> <li>• Wysa realized the good feedback</li> </ul> |

#### Test case 8

|                       |  |  |
|-----------------------|--|--|
| Test case ID          | E8   |  |
| Test case description | User wants to do reframe thoughts (once)   |  |
| Steps                 | Wysa   | User   |
| 1                     |  | “I want to reframe thoughts to feel better as an Asian teen” |
| 2                     | Wysa guides user through reframe thoughts  | User follows it and reframe their thoughts                   |
| 3                     |  | “The exercise was unhelpful”                                 |
| Actual output         | <ul style="list-style-type: none"> <li>• Wysa was able to guide user through the reframing thoughts process</li> <li>• Wysa did not offer personalization</li> <li>• Wysa realized the unhelpful feedback</li> </ul> |  |

#### Test case 9

|                       |  |  |
|-----------------------|--|--|
| Test case ID          | E9   |  |
| Test case description | User wants to do strength exercises (once) |  |
| Steps                 | Wysa                                       | User   |
| 1                     |  | “Give me some strength exercises for Asian elders” |

|               |  |                                    |
|---------------|--|------------------------------------|
| 2             | Wysa offers some physical exercises  |                                    |
| 3             |  | User chooses one of the exercises  |
| 4             | Wysa guides user through the exercise  | User follows the exercises         |
| 5             |  | User asks what exercises been done |
| Actual output | <ul style="list-style-type: none"> <li>• Wysa offers physical exercises</li> <li>• Wysa offers no personalization</li> <li>• Wysa ignored the question for what exercises been done</li> </ul> |                                    |

#### Test case 10

|                       |   |   |
|-----------------------|---|---|
| Test case ID          | E10   |   |
| Test case description | User wants to set exercise routines for physical strength exercises   |   |
| Steps                 | Wysa  | User  |
| 1                     |   | “I want to set some strength exercise routines as an Asian elder” |
| 2                     | Wysa offers some physical exercises   |   |
| Actual output         | <ul style="list-style-type: none"> <li>• Wysa mistaken request for physical exercise routine as doing exercise</li> </ul> |   |

#### Test case 11

|                       |   |  |
|-----------------------|---|--|
| Test case ID          | E11   |  |
| Test case description | User wants to give good feedback to physical strength exercises |  |
| Steps                 | Wysa  | User   |
| 1                     |   | “Give me some strength exercises for Asian elders” |

|               |  |                                    |
|---------------|--|------------------------------------|
| 2             | Wysa offers some personalized strength exercises   |                                    |
| 3             |  | User chooses one of the exercises  |
| 4             | Wysa guides user through the exercise  | User follows the exercises         |
| 5             |  | "This exercise was helpful for me" |
| Actual output | <ul style="list-style-type: none"> <li>• Wysa was able to guide through the exercises</li> <li>• Wysa was not able to offer personalized exercises</li> <li>• Wysa realized the feedback was good</li> </ul> |                                    |

### Test case 12

User wants to give bad feedback for physical strength exercise

|                       |   |  |
|-----------------------|---|--|
| Test case ID          | E12   |  |
| Test case description | User wants to give bad feedback to mindfulness exercises  |  |
| Steps                 | Wysa  | User   |
| 1                     |   | "Give me some strength exercises for Asian elders" |
| 2                     | Wysa offers some physical exercises   |  |
| 3                     |   | User chooses one of the exercises                  |
| 4                     | Wysa guides user through the exercise   | User follows the exercises                         |
| 5                     |   | "This exercise was not helpful for me"             |
| 6                     | Wysa acknowledged the feedback  |  |
| Actual output         | <ul style="list-style-type: none"> <li>• Wysa was able to guide the user through the exercise</li> <li>• Wysa did not offer physical exercises personalization</li> <li>• Wysa acknowledged the feedback</li> </ul> |  |

### Test case 13

|                       |  |  |
|-----------------------|--|--|
| Test case ID          | E13  |  |
| Test case description | User wants to do durability exercises (once)   |  |
| Steps                 | Wysa   | User   |
| 1                     |  | “Give me some durability exercises for unemployed adult” |
| 2                     | Wysa offers some physical exercises  |  |
| 3                     |  | User chooses one of the exercises                        |
| 4                     | Wysa guides user through the exercise  | User follows the exercises                               |
| 5                     |  | User asks what exercises been done                       |
| Actual output         | <ul style="list-style-type: none"> <li>• Wysa only offered same exercises as strength exercises</li> <li>• Wysa offered no personalization</li> <li>• Wysa had no knowledge of what exercises been done</li> </ul> |  |

### Test case 14:

|                       |   |   |
|-----------------------|---|---|
| Test case ID          | E14   |   |
| Test case description | User wants to set exercise routines for mindfulness exercises   |   |
| Steps                 | Wysa  | User  |
| 1                     |   | “I want to set routine to do mindfulness every Monday as an Latino adult” |
| 2                     | Wysa offers mindfulness exercises instead   |   |
| Actual output         | <ul style="list-style-type: none"> <li>• Wysa was not able to offer routine and mistaken for wanting to do exercises</li> </ul> |   |

Test case 15:

User wants to give good feedback for durability exercise

|                       |   |   |
|-----------------------|---|---|
| Test case ID          | E15   |   |
| Test case description | User wants to give good feedback to mindfulness exercises   |   |
| Steps                 | Wysa  | User  |
| 1                     |   | "Give me some mindfulness exercises for Asian elders" |
| 2                     | Wysa offers some personalized mindfulness exercises   |   |
| 3                     |   | User chooses one of the exercises                     |
| 4                     | Wysa guides user through the exercise   | User follows the exercises                            |
| 5                     |   | "This exercise was helpful for me"                    |
| Actual output         | <ul style="list-style-type: none"><li>• Wysa was able to acknowledge that the exercises was helpful</li></ul> |   |

Test case 16:

|                       |   |  |
|-----------------------|---|--|
| Test case ID          | E16   |  |
| Test case description | User wants to give bad feedback to durability exercises |  |
| Steps                 | Wysa  | User   |
| 1                     |   | "Give me some durability exercises for Asian elders" |
| 2                     | Wysa offers some personalized durability exercises      |  |
| 3                     |   | User chooses one of the exercises                    |

|               |   |  |
|---------------|---|--|
| 4             | Wysa guides user through the exercise   | User follows the exercises             |
| 5             |   | “This exercise was not helpful for me” |
| Actual output | <ul style="list-style-type: none"> <li>Wysa was able to acknowledge that the exercises was not helpful</li> </ul> |  |

Test case 17:

|                       |   |   |
|-----------------------|---|---|
| Test case ID          | E17   |   |
| Test case description | User wants to give feedback to physical exercises   |   |
| Steps                 | Wysa  | User  |
| 1                     |   | “Tôi muốn tập thể dục” (I want to exercise but in Vietnamese) |
| 2                     | <p>“You can explore the slider below for quick access to some of the features.”</p> <p>“You can also check your self-care kit for something that might help you”</p> <p>Show many exercises</p> |   |
| Actual output         | <ul style="list-style-type: none"> <li>Wysa was not able to detect different languages</li> </ul>   |   |

Test case 18

|                       |  |      |
|-----------------------|--|------|
| Test case ID          | E18  |      |
| Test case description | User wants to change exercises to physical exercises |      |
| Steps                 | Wysa   | User |

|               |  |                                   |
|---------------|--|-----------------------------------|
| 1             |  | "Let's get mindfulness"           |
| 2             | Wysa offers some mindfulness exercises   |                                   |
| 3             |  | User chooses one of the exercises |
| 4             | Wysa guides user through the exercise  |                                   |
| 5             |  | "Let's get active"                |
| 6             | Wysa offers some physical exercises  |                                   |
| 7             |  | "Let's get active"                |
| 8             | Wysa offers some physical exercises  |                                   |
| Actual output | <ul style="list-style-type: none"> <li>Wysa was able to change according to the request user made</li> </ul> |                                   |

#### Augmented test case

For AI testing augmented test cases, we will use test case 1 as the base case

| ID | Augment Profile    | Augment Result  | Wysa able to recognize request for mindfulness exercise | Wysa offers exercise personalization |
|----|--------------------|---|---|--------------------------------------|
| 1  | random word swap   | Give some me mindfulness exercises for Asian elders             | Yes   | No                                   |
| 2  | random word delete | me some exercises for Asian elders                              | Yes   | No                                   |
| 3  | random word insert | Give heedfulness me some mindfulness exercises for Asian elders | Yes   | No                                   |

|   |                       |   |                                       |    |
|---|-----------------------|---|---------------------------------------|----|
| 4 | synonym augmentation  | Afford me some heedfulness exercises for Asian elders | No (offer physical exercises instead) | No |
| 5 | ocr augmentation      | Give me some mindfolne88 exercises for Asian eldeks   | Yes                                   | No |
| 6 | keyboard augmentation | Give me some mindfulness exe3dises for Asian #lders   | Yes                                   | No |
| 7 | random char insert    | wGive me some mindfulness exercises for A2sian elders | Yes                                   | No |
| 8 | random char swap      | Gvie me some mindfulness exercises for Asian eldres   | Yes                                   | No |
| 9 | random char delete    | Give me som mindfulness exercises for Asian elers     | Yes                                   | No |

### Invalid Topics Function

Test case 1:

|              |                           |                                |
|--------------|---------------------------|--------------------------------|
| Test case ID | I1                        |                                |
| Input        | Topic: Cooking            | Type: Question                 |
| Steps        | Wysa                      | User                           |
| 1            | Ask how the user is doing |                                |
| 2            |                           | How do you know if a burger is |

|               |   |              |
|---------------|---|--------------|
|               |   | done or not? |
| 3             | Let's talk about that. Go on...   |              |
| Actual output | Wysa was not able to realize that the question was not related to mental health |              |
| Result        | Completely illogical  |              |

Test case 2:

|               |   |  |
|---------------|---|--|
| Test case ID  | I2  |  |
| Input         | Topic: Cooking  | Type: Statement                                  |
| Steps         | Wysa  | User   |
| 1             | Ask how the user is doing   |  |
| 2             |   | Chicken is safe to eat at 165 degrees Fahrenheit |
| 3             | Let's talk about that. Go on...   |  |
| Actual output | Wysa was not able to realize that the question was not related to mental health |  |
| Result        | Not completely illogical but not correct  |  |

Test case 3:

|              |                                 |  |
|--------------|---------------------------------|--|
| Test case ID | I3                              |  |
| Input        | Topic: Stock Market             | Type: Question                                 |
| Steps        | Wysa                            | User   |
| 1            | Ask how the user is doing       |  |
| 2            |                                 | What is Downey's stock looking like right now? |
| 3            | Let's talk about that. Go on... |  |

|               |   |
|---------------|---|
| Actual output | Wysa was not able to realize that the question was not related to mental health |
| Result        | Completely illogical  |

Test case 4:

|               |   |   |
|---------------|---|---|
| Test case ID  | I4  |   |
| Input         | Topic: Stock Market   | Type: Statement   |
| Steps         | Wysa  | User  |
| 1             | Ask how the user is doing   |   |
| 2             |   | The general rule of thumb in the stock market is to buy low and sell high |
| 3             | What's got you feeling this way, [user]?  |   |
| Actual output | Wysa was not able to realize that the question was not related to mental health |   |
| Result        | Completely illogical  |   |

Test case 5:

|              |  |   |
|--------------|--|---|
| Test case ID | I5   |   |
| Input        | Topic: Sports  | Type: Question                            |
| Steps        | Wysa   | User                                      |
| 1            | Ask how the user is doing  |   |
| 2            |  | Why were the Washington Redskins renamed? |
| 3            | Scientists say that positive emotions increase one's awareness and encourage novel |   |

|               |  |  |
|---------------|--|--|
|               | thoughts and actions. This could be a fantastic day to start something creative. |  |
| 4             | What's the main focus for today?   |  |
| Actual output | Wysa was not able to realize that the question was not related to mental health  |  |
| Result        | Completely illogical   |  |

Test case 6:

|               |   |  |
|---------------|---|--|
| Test case ID  | I6  |  |
| Input         | Topic: Sports   | Type: Statement  |
| Steps         | Wysa  | User   |
| 1             | Ask how the user is doing   |  |
| 2             |   | American football is not the same as football in other countries |
| 3             | Scientists say that positive emotions increase one's awareness and encourage novel thoughts and actions. This could be a fantastic day to start something creative. |  |
| 4             | What's the main focus for today?  |  |
| Actual output | Wysa was not able to realize that the question was not related to mental health   |  |
| Result        | Completely illogical  |  |

Test case 7:

|              |    |
|--------------|----|
| Test case ID | I7 |
|--------------|----|

|               |   |                         |
|---------------|---|-------------------------|
| Input         | Topic: Weather  | Type: Question          |
| Steps         | Wysa  | User                    |
| 1             | Ask how the user is doing   |                         |
| 2             |   | Will it rain next week? |
| 3             | Oh! Go on, [user]. I'm listening.   |                         |
| Actual output | Wysa was not able to realize that the question was not related to mental health |                         |
| Result        | Completely illogical  |                         |

Test case 8:

|               |   |  |
|---------------|---|--|
| Test case ID  | I8  |  |
| Input         | Topic: Weather  | Type: Statement  |
| Steps         | Wysa  | User   |
| 1             | Ask how the user is doing   |  |
| 2             |   | California experienced turbulent weather changes in 2023 |
| 3             | Is that still bothering you, [user]?  |  |
| Actual output | Wysa was not able to realize that the question was not related to mental health |  |
| Result        | Logical but incorrect response  |  |

Test case 9:

|              |             |                |
|--------------|-------------|----------------|
| Test case ID | I9          |                |
| Input        | Topic: Math | Type: Question |
| Steps        | Wysa        | User           |

|               |   |                              |
|---------------|---|------------------------------|
| 1             | Ask how the user is doing   |                              |
| 2             |   | What is the derivative of 8? |
| 3             | Let's talk about that. Go on...   |                              |
| Actual output | Wysa was not able to realize that the question was not related to mental health |                              |
| Result        | Completely illogical  |                              |

Test case 10:

|               |  |                      |
|---------------|--|----------------------|
| Test case ID  | I10  |                      |
| Input         | Topic: Math  | Type: Statement      |
| Steps         | Wysa   | User                 |
| 1             | Ask how the user is doing  |                      |
| 2             |  | 4 is the square of 2 |
| 3             | Okay...  |                      |
| 4             | [user], if you were to try and find your mood now, what kind of an 'okay' are you feeling right now? |                      |
| Actual output | Wysa was not able to realize that the question was not related to mental health                      |                      |
| Result        | Completely illogical   |                      |

Augmented test case

Test case 1 is used as the example

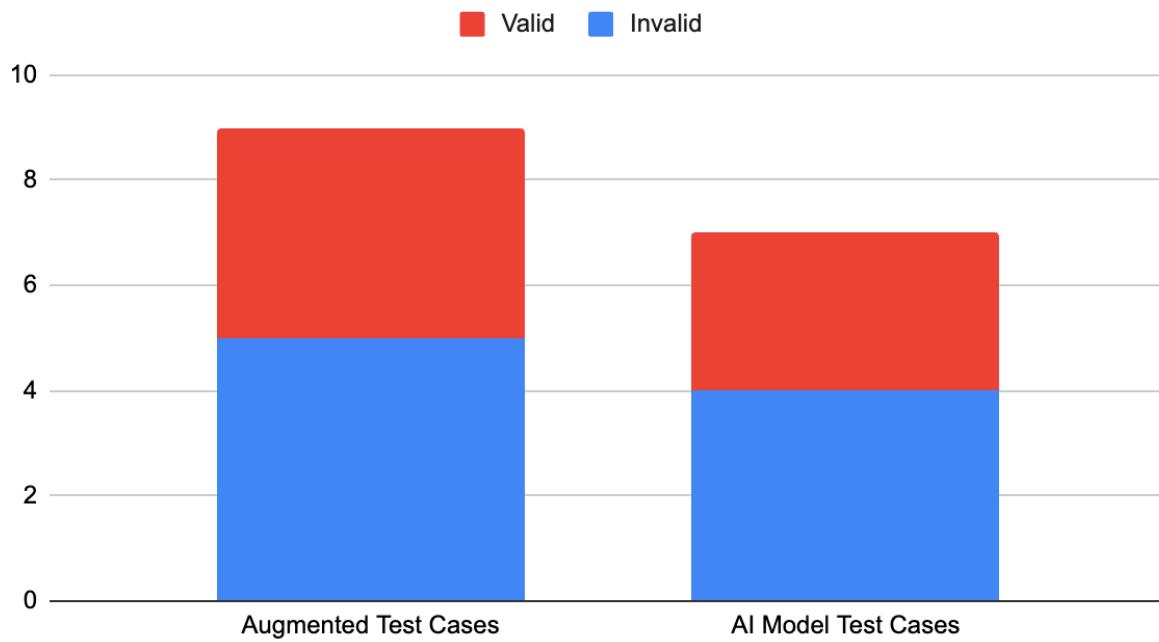
| ID | Augmentation       | Augmentation result                         | Wysa is able to detect that it is not a valid topic |
|----|--------------------|---|---|
| 1  | random word swap   | How do or know if a burger is done you not? | No  |
| 2  | random word delete | do you if a burger is not?                  | Yes but not really                                  |

|   |                       |  |    |
|---|-----------------------|--|----|
| 3 | random word insert    | How do you know if a burger is done bonk or not? | No |
| 4 | synonym augmentation  | How serve you know if a burger is done or non?   | No |
| 5 | ocr augmentation      | How do you know if a borger i8 done or n0t?      | No |
| 6 | keyboard augmentation | How do you knkw if a b*rger is doBe or not?      | No |
| 7 | random char insert    | How do you knobw if a burDger is d4one or not?   | No |
| 8 | random char swap      | How do you nkow if a bruger is odne or not?      | No |
| 9 | random char delete    | How do you knw if a burge is don or not?         | No |

## Section 3.3 Test case analysis (statistics)

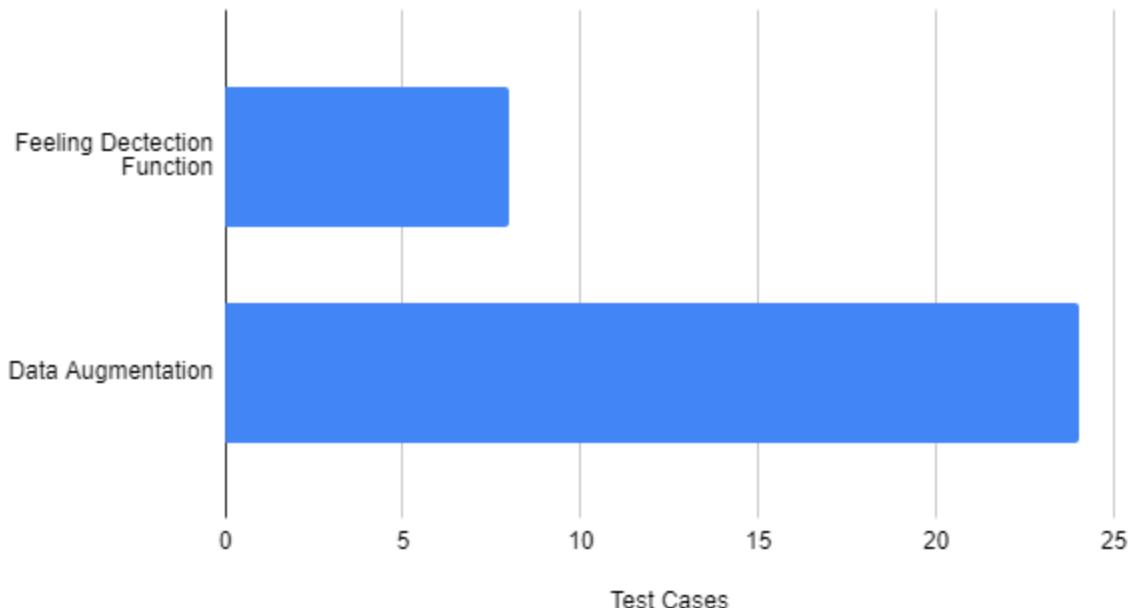
Mental Illness Function

### Mental Illness Test Case Analysis



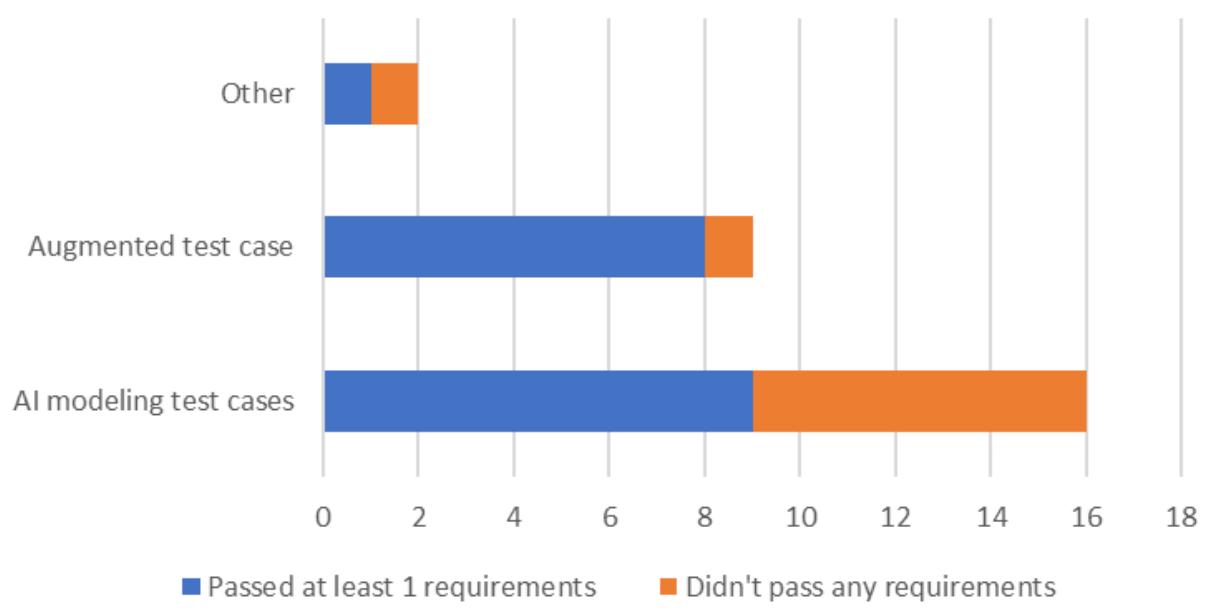
Feeling Detection Function

### Overall Test Cases For Feeling Detection Function



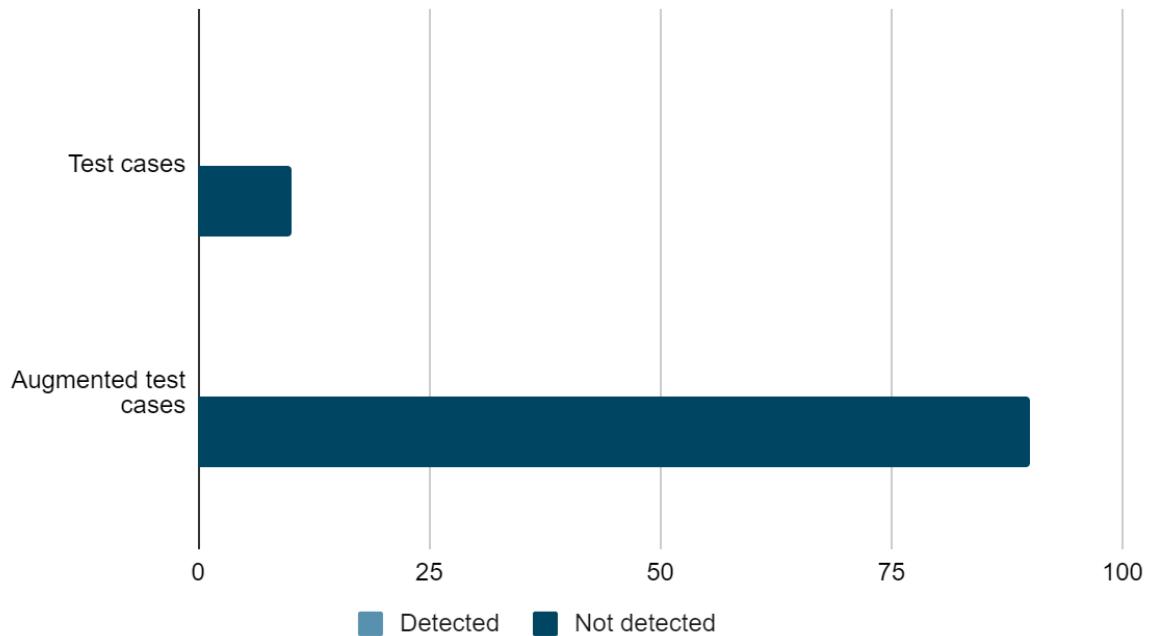
### Exercise function

#### Exercise function test report



## Invalid Topics Function

### Invalid Topic Detection



## Section 4 – AI Function Test Results and Test Criteria

### Section 4.1 AI test model-based test results analysis test complexity, and test result statistics

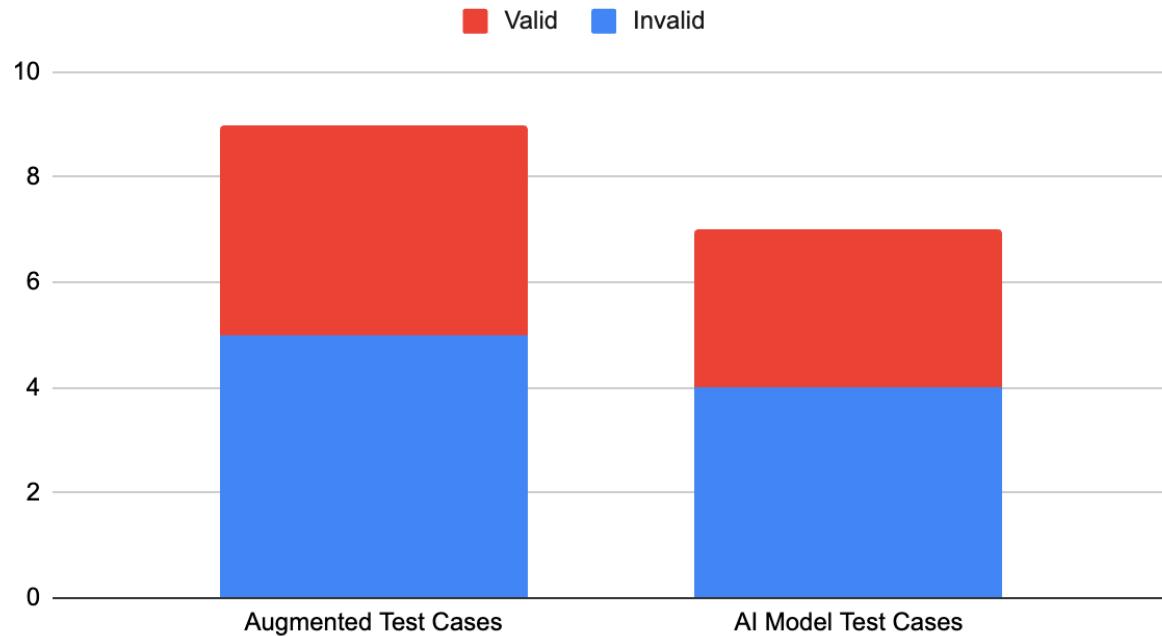
#### Mental Illness Function

Test case number:

- There are 8 test cases, 9 augmented test cases
- The test complexity is 17
- Each test covers:
  - Input
    - Mental Issues
      - Disorder
        - Anxiety, Sleep
      - Illness
        - Depression, Stress
    - Causes

- Trauma
  - Events
- Negative
  - Thoughts, Feelings
- Context
  - Age
    - Group
      - Teen, Adult, Elder
  - Occupation
    - Student
      - High School, College
    - Employee
      - Full Time, Part Time, Unemployed
- Conversation Format
- Memory
  - Short term memory
  - Long term memory
- Different languages
- Text Data Augmentation

## Mental Illness Test Case Analysis

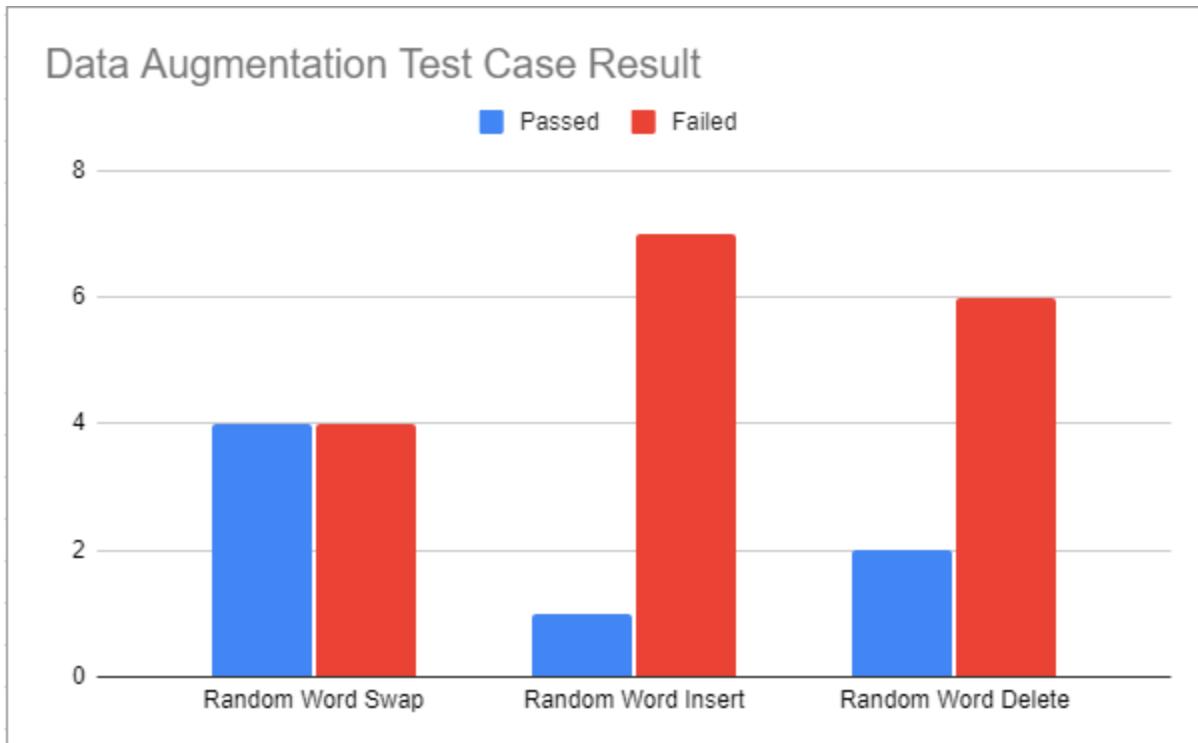


## Feeling Detection Function

Test case number:

- There are 8 test cases, 24 augmented test cases
- The test complexity is 32
- Each test covers the following topics:
  - Input:
    - Good Feelings: Happy, Excited
    - Sad Feelings: Sad, Anxious, Stress
  - Context:
    - Age group: Teen, Adult, Elder
    - Occupation: Employee, Unemployed
    - Language: English, and Non-English
  - Data Augmentation:
    - Random word swap
    - Random word insert
    - Random word delete
  - Conversation Format





## Exercise function

### Test complexity:

We have 27 AI test model-based test cases total:

- 16 AI-modeling test cases
- 9 augmented test cases
- 1 test case to test different language
- 1 test case to test conversation format

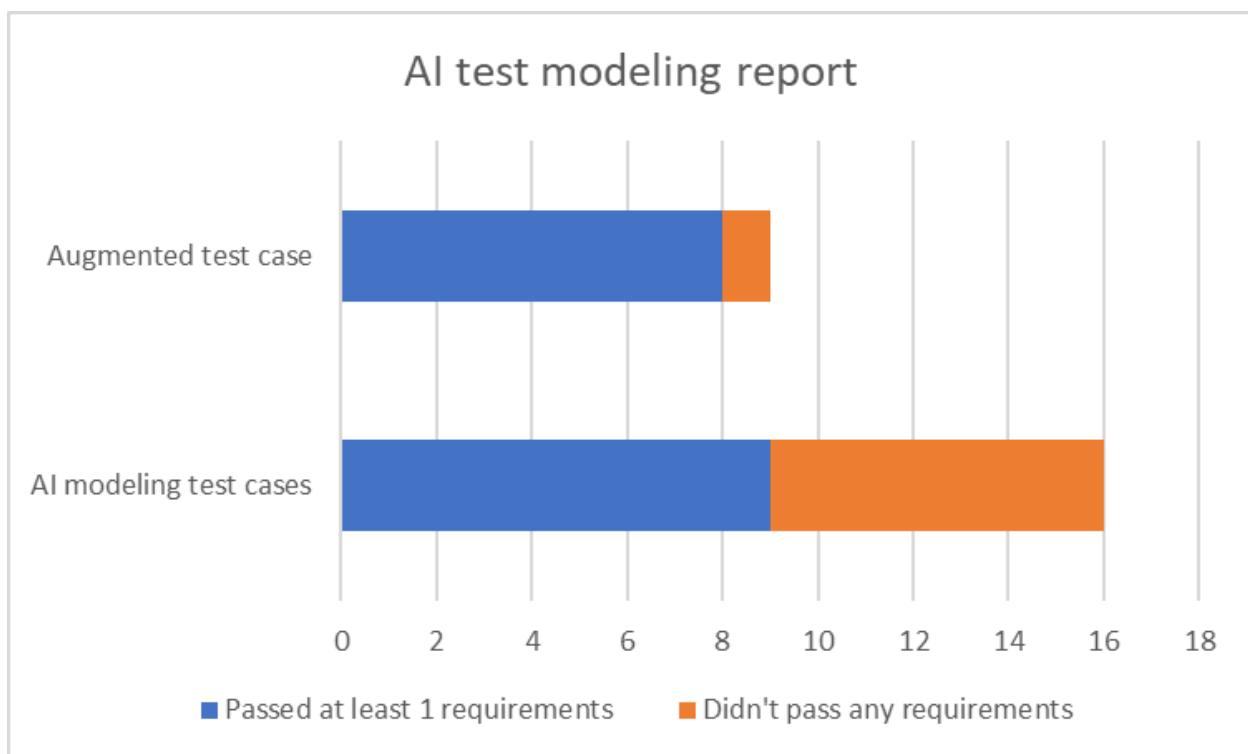
For exercise input, there is at least one test case cover each type of input

- Exercises
  - Mental
    - Reframe thoughts
    - Mindfulness
  - Physical
    - Strength
    - Endurance
- Conversation
  - Do exercise
    - One time
    - Set exercise schedule

- Feedback
  - Good
  - Bad

Data augmentation for test case 1:

- random word swap
- random word delete
- random word insert
- synonym augmentation
- ocr augmentation
- keyboard augmentation
- random char insert
- random char swap
- random char delete



- For AI test modeling report
  - Augmented test case passed 8 out of 9, 88.889% passed
  - AI modeling test cases passed 9 out of 16, 56.25% passed

## Invalid Topics Function

Test case amount:

- There are a total 10 test case with 9 augmentations each for a total 90 augmented test cases

- Test complexity is 100
- Test coverage:
  - Input:
    - Question: Cooking, Stock market, Sports, Weather, Math
    - Statement: Cooking, Stock market, Sports, Weather, Math
  - Context:
    - Age: Teen, Adult, Elderly
    - Language: English, Non-English
  - Augmentation:
    - random word swap
    - random word delete
    - random word insert
    - synonym augmentation
    - ocr augmentation
    - keyboard augmentation
    - random char insert
    - random char swap
    - random char delete

## Section 4.2 AI Function bug analysis

### Mental Illness Function

- Wysa is not able to understand different languages and can only respond in english
- When User doesn't follow Wysa's conversation format, Wysa will provide irrelevant suggestions/recommendations or will not understand the User's needs
- Wysa was able to remember short term mental health and causes inputs and long term mental health and causes inputs
- Wysa fails 55.55% of the text data augmentation test cases
- Wysa fails 57.143% of the AI Model test cases

### Feeling Detection Function:

Age Group: During our testing, we conducted an assessment to determine whether Wysa, the emotional support chatbot, could accurately identify the age group of its users. We tested a variety of age groups to determine Wysa's proficiency in identifying different age ranges. However, our results showed that Wysa is unable to detect the age group of its users with precision. This means that Wysa's ability to provide support and guidance is not affected by the age group of the user.

Language: During our testing, we tested Wysa to see if it is able to communicate with different languages to its users. However, our results showed that Wysa is not capable

of comprehending languages other than English. This means that Wysa's language proficiency is limited to English, and it cannot provide effective emotional support to users who communicate in other languages.

**Occupation:** As part of our testing process, we evaluated Wysa's ability to identify the occupation of its users. We conducted several tests to assess whether Wysa could detect and respond appropriately to users with different occupations. However, our results showed that Wysa is unable to identify the occupation of its users accurately. Wysa's ability to provide emotional support and guidance is not influenced by the user's profession.

## Exercise function

- Based on our testing analysis, Wysa exercise function is very limited, the only thing it offers is guiding users through either mental or physical exercises.
- Wysa, at the time of testing, does not offer exercise personalization based on different age groups, occupation, etc.
- Feedback on Wysa exercise is also limited
- Wysa did not pass the different language test
- Wysa likely to offers mindfulness exercise if the word 'mindfulness' is in the sentence
- Wysa likely to offers physical exercise if the word 'exercise' or 'active' is in the sentence

## Invalid Topics Function

- Analyzing the test results dictate that Wysa is unable to detect if the conversation is actually related to mental health. This is fundamentally an issue because this means the app will interpret anything and everything as related to the user talking about their mental state.
- The app is also incompatible with languages other than English. This is obviously inconvenient for people that are unable to speak or understand English.

## Section 4.3 AI function test quality assessment

### Mental Illness Function

- This test for detecting the User's mental illness and providing mental health support covers the app's capability to test the following
  - Different types of mental illnesses (disorders and illnesses)
    - Depression and Stress are being tested
    - Sleep and Anxiety are being tested
  - Different types of causes to their mental illnesses
    - Traumatic events
    - Negative thoughts and feelings
- In addition to the app's main functionality, we tested their capability to understand incomprehensible text using the AI test tool's data augmentation
- We tested Wysa's ability to understand or respond in a different language
- Tested different conversation formats that tested Wysa's ability to adjust mid conversation
- We also tested Wysa's capability to have a short term and long term memory

### Exercise function

- The test suite for exercise function is still very basic but we believe that it covers enough depth for an exercise chat app, such as
  - Cover different type of exercises, such as mental and physical
    - In mental we asks for mindfulness and reframe
    - In physical we ask for strength and durability
  - Cover different type of others
    - Such as setting up schedule
    - Give feedback
  - Cover user's occupation, age, group, etc
- We have not testing advanced topics such as calorie tracking, set custom timer, etc
- However, this test suite is enough to test Wysa's exercise functionality since Wysa was not able to support all of the test suite requirements

### Invalid Topics Function

- The test results show that while the app is capable of detecting mental states relatively well, it fails to detect topics that are not related to mental health correctly, such as but not limited to:

- Cooking
- Stock market trends
- Sports
- Weather
- Math
- While there is an infinite amount of alternate topics that could also be used for testing, we opted to test a limited topic list that would help us infer the overall pass rate for the app.

**Deliverable 2 - AI Testing Tool Report**  
**CMPE 187 - Software Quality Engineering**  
**San Jose State University**

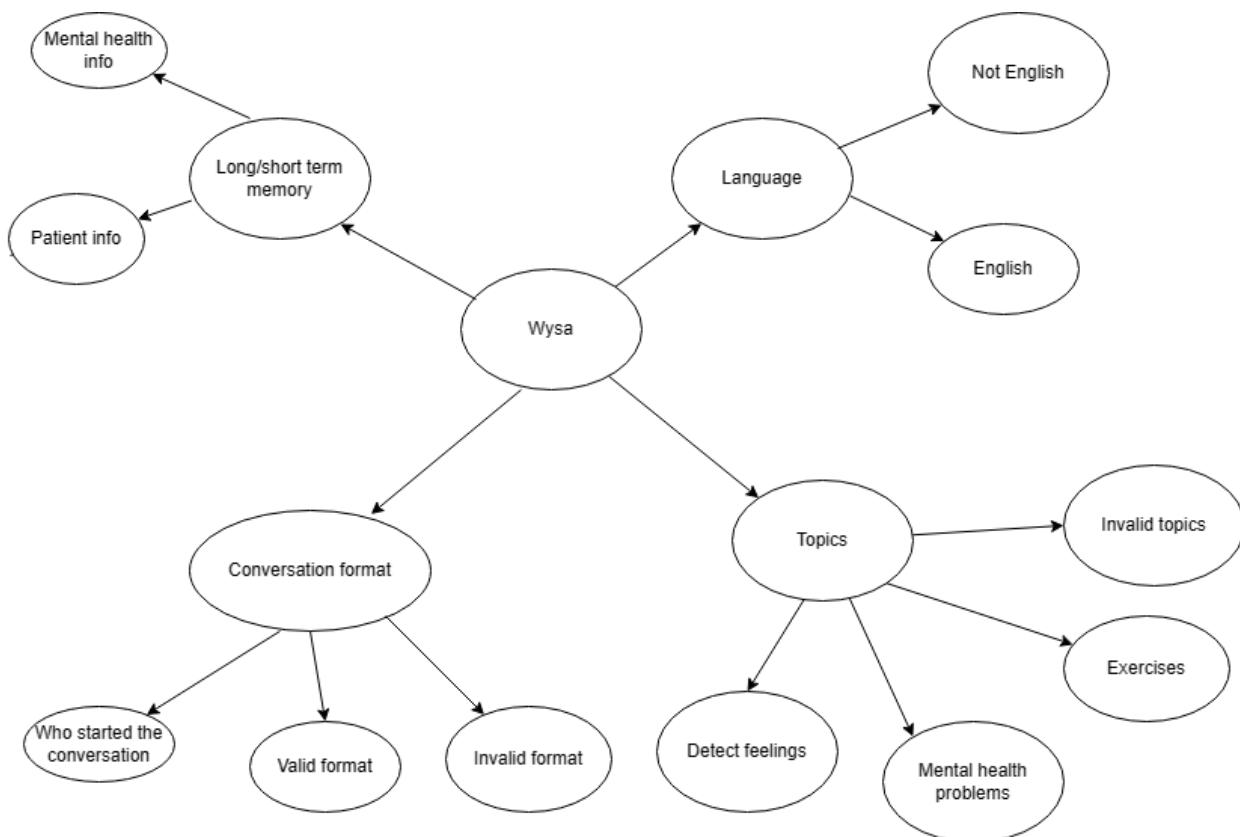
**Group 2**

Team members:  
Khai Nguyen  
Faith Deanon  
Kevin Campbell  
Long Nguyen

# Section 1 – AI Function Test Requirement Analysis

## 1. AI function test requirements

- Wysa is a chatbot that offers mental health counseling through chatting. In addition, it also has functionalities such as guiding through exercises.
- We decided to narrow down Wysa's core functionalities to perform our AI testing, those are:
  - Detect feelings, whether or not Wysa can detect user's feelings through chatting
  - Mental health problems, can Wysa detect mental health problems just by chatting
  - Exercises, what type of exercises can Wysa offer as well as functionalities like giving feedback
  - Invalid topics, we will also test invalid topics to see whether or not Wysa can handle those.



## 2. AI function test requirement modeling

- To test functionality above, we will use an AI testing method consisting of specifying input, context and output for each function above.

- Input specify type of input depends on the functions. Context specifies type of user and output categories type of output Wysa offers.

## Section 2 – AI Test Modeling for Selected AI Features

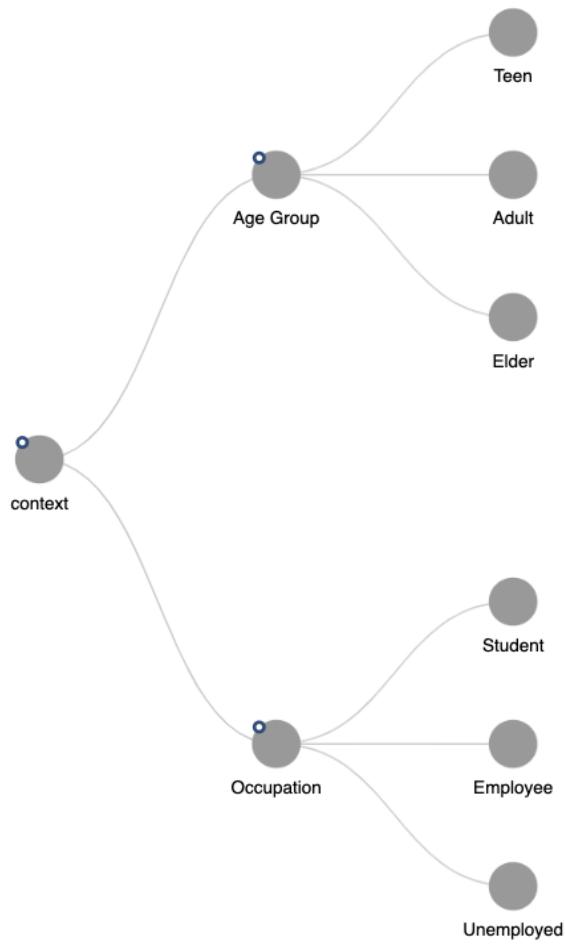
AI testing tool was used in order to produce input, context, output and classification decision tables

1. Context modeling for each selected AI-powered function/feature
2. AI-powered function input classifications
3. AI-powered function output/event/action classifications
4. AI-powered function classification decision tables

### Mental Illness - Faith

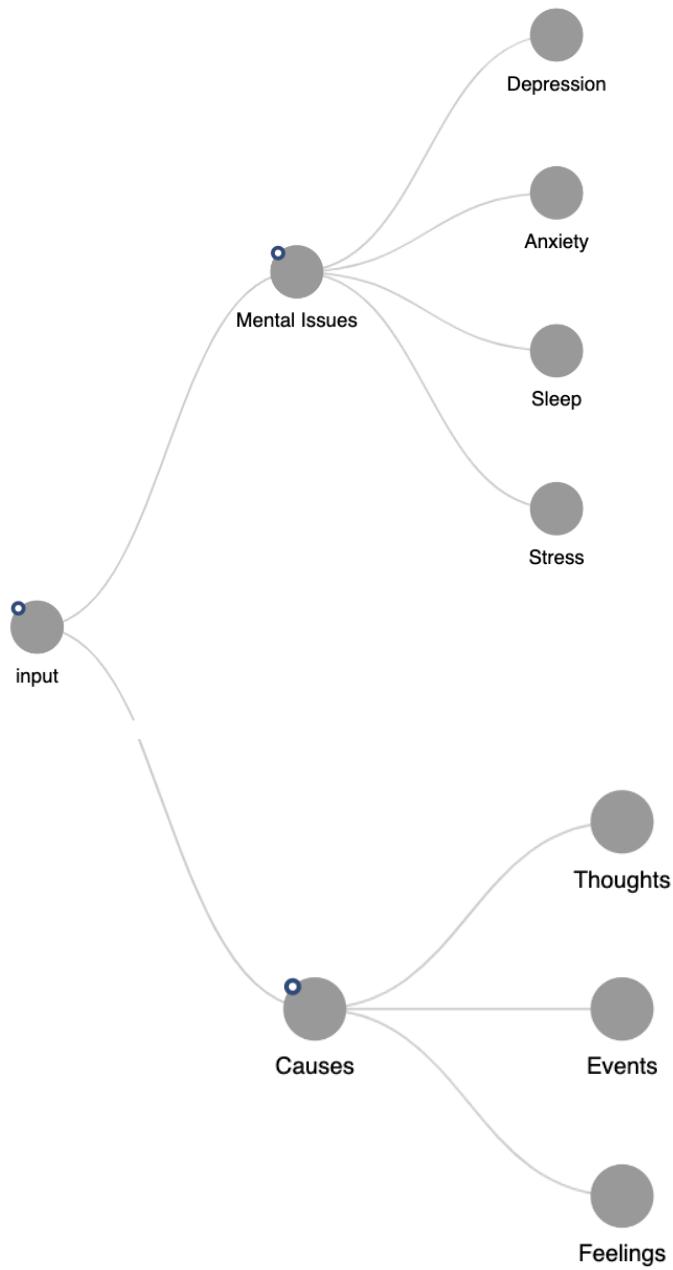
#### 1. AI-powered Function Context Model Classification

This context model represents users based on their age group and occupation. A User can either be a teen, adult, or elder. A User's occupation can be Student, Employee, or Unemployed. Wysa will be tested based on various contextual factors.



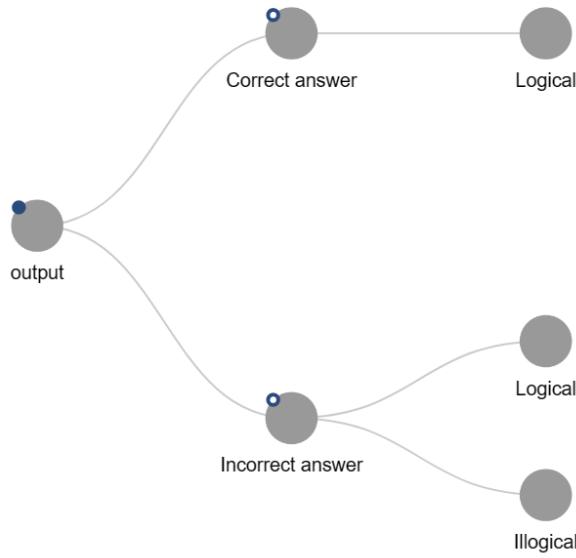
## 2. AI Powered Function Input Classification

Wysa will be tested based on two different types of user inputs: Mental Issues and Causes. The app will test various combinations of different types of mental issues and types of causes. The types of mental issues that will be considered are depression, anxiety, sleep, and stress. The types of causes that will be tested are thoughts, events, and feelings.



### 3. AI-powered function output/event/action classifications

This output model consists of two output classifications: correct answer, incorrect answer. A correct answer output will be logical while an incorrect answer can either be logical or illogical. This output model is shared with the Exercise function and is explained in length in the Exercise Function's AI-powered function output/event/action classifications.



#### 4. AI-powered function classification decision table

##### Input Decision Table

The AI test tool created an input decision table based on the input model. It generated a total of 12 combinations of user inputs.

| @id | Causes   | Mental Issues |
|-----|----------|---------------|
| 1   | Thoughts | Depression    |
| 2   | Thoughts | Stress        |
| 3   | Events   | Stress        |
| 4   | Feelings | Stress        |
| 5   | Events   | Depression    |
| 6   | Feelings | Depression    |
| 7   | Thoughts | Anxiety       |
| 8   | Events   | Anxiety       |
| 9   | Feelings | Anxiety       |
| 10  | Thoughts | Sleep         |
| 11  | Events   | Sleep         |
| 12  | Feelings | Sleep         |

### Context Decision Table

The AI tool created a context decision table based on the context model. It generated test cases based on various context classifications. It created 9 combinations of context classifications with age groups and occupations.

| @id | Age Group | Occupation |
|-----|-----------|------------|
| 1   | Teen      | Student    |
| 2   | Teen      | Employee   |
| 3   | Teen      | Unemployed |
| 4   | Adult     | Student    |
| 5   | Adult     | Employee   |
| 6   | Adult     | Unemployed |
| 7   | Elder     | Student    |
| 8   | Elder     | Employee   |
| 9   | Elder     | Unemployed |

### Output Decision Table

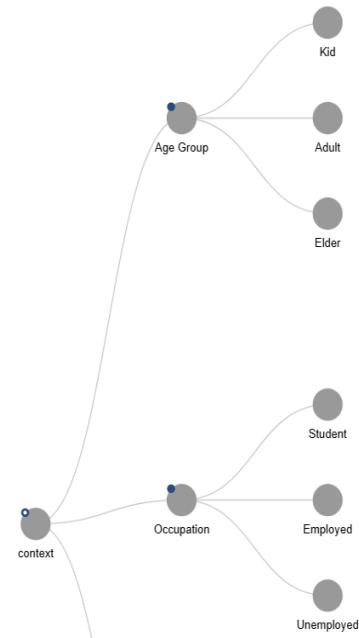
Based on the output model, the AI tool created an output decision table. It created 12 outputs because there are 12 inputs. However, each test case has outputs that are undefined. This means that each output has to be filled out manually.

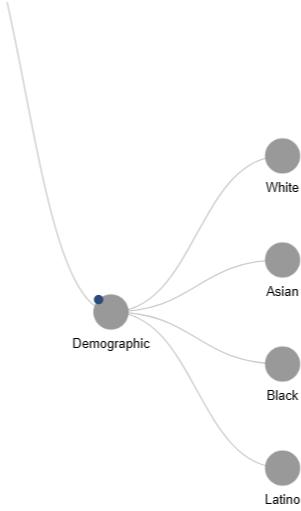
| @id | Invalid   | Valid     |
|-----|-----------|-----------|
| 1   | undefined | undefined |
| 2   | undefined | undefined |
| 3   | undefined | undefined |
| 4   | undefined | undefined |
| 5   | undefined | undefined |
| 6   | undefined | undefined |
| 7   | undefined | undefined |
| 8   | undefined | undefined |
| 9   | undefined | undefined |
| 10  | undefined | undefined |
| 11  | undefined | undefined |
| 12  | undefined | undefined |

## Exercises - Khai

Aside from mental health counseling, Wysa chatbot can also guide users through different exercises. This is to test whether or not Wysa is good or versatile in conversing with users through exercises

### 1. Context modeling for each selected AI-powered function/feature

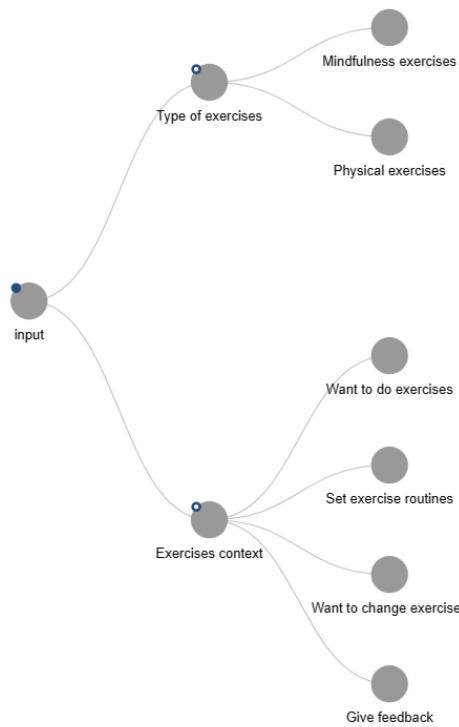




*Figure 3.1.1 and 3.1.2, context modeling for Exercises function*

The context grouped users into 3 different categories, age group, occupation and demographic. For age groups, the user is divided to kid, adult and elder. Occupation is further divided into student, employed and unemployed. In demographic, there are four categories: white, asian, black and latino.

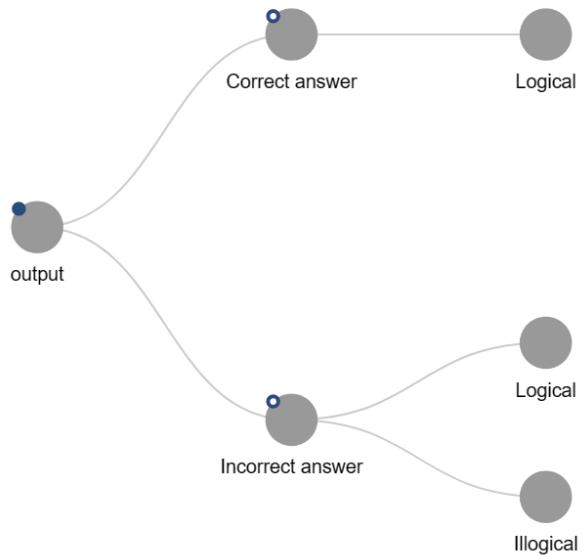
## 2. AI-powered function input classifications



*Figure 3.2.1, input modeling for Exercises function*

The input for the exercise function is divided into two categories. First is the type of exercises and second is the context. Type of exercise contains mindfulness exercises and physical exercises. In exercises context, it is divided into want to do exercises, set exercises routines, want to change exercises and give feedback.

### 3. AI-powered function output/event/action classifications



*Figure 3.3.1, output modeling for Exercises function*

For the output, there are correct answers and incorrect answers. If the answer is correct, then it has to be logical. If it is not correct, it can still be logical, but not related to the current context, or it can simply be illogical.

#### 4. AI-powered function classification decision tables

| <i>id</i> | <i>Age Group</i> | <i>Demographic</i> | <i>Occupation</i> |
|-----------|------------------|--------------------|-------------------|
| 1         | Kid              | White              |                   |
| 2         | Kid              | Asian              | Student           |
| 3         | Kid              | Black              | Student           |
| 4         | Kid              | Latino             | Student           |
| 5         | Kid              | White              | Student           |
| 6         | Kid              | Asian              | Employed          |
| 7         | Kid              | Black              | Employed          |
| 8         | Kid              | Latino             | Employed          |
| 9         | Kid              | White              | Employed          |
| 10        | Kid              | Asian              | Unemployed        |

*Figure 3.4.1 Classification decision tables top view*

| id | Context   | Semantic | id | Exercises context        | Type of exercises     |
|----|-----------|----------|----|--------------------------|-----------------------|
| 1  | Correct   | Logical  | 1  | Want to do exercises     | Mindfulness exercises |
| 2  | Correct   | Logical  | 2  | Set exercise routines    | Mindfulness exercises |
| 3  | Correct   | Logical  | 3  | Want to change exercises | Mindfulness exercises |
| 4  | Correct   | Logical  | 4  | Give feedback            | Mindfulness exercises |
| 5  | Correct   | Logical  | 5  | Want to do exercises     | Physical exercises    |
| 6  | Correct   | Logical  | 6  | Set exercise routines    | Physical exercises    |
| 7  | Correct   | Logical  | 7  | Want to change exercises | Physical exercises    |
| 8  | undefined | Logical  | 8  | Give feedback            | Physical exercises    |

*Figure 3.4.2, Classification decision table rear view*

| <b>id</b> | <b>Exercises context</b> | <b>Type of exercises</b> |
|-----------|--------------------------|--------------------------|
| 1         | Want to do exercises     | Mindfulness exercises    |
| 2         | Want to change exercises | Mindfulness exercises    |
| 3         | Give feedback            | Mindfulness exercises    |
| 4         | Want to do exercises     | Physical exercises       |
| 5         | Want to change exercises | Physical exercises       |
| 6         | Give feedback            | Physical exercises       |

Figure 3.4.3, Classification decision table rear view different angle

The classification decision table is based on input, context and output specified above.

## Feeling Detection Function - Long

### Input model

The input classification spanning tree for feeling detection encompasses a comprehensive framework that categorizes emotions into two major cases and five minor cases. The first major case is centered around negative or unpleasant feelings, which encompasses three minor cases including stress, sadness, and anxiety. These emotions are typically associated with distress, discomfort, or unease, and may manifest in various ways in different individuals.

The second major case in the classification spanning tree pertains to positive or pleasant feelings, including emotions such as happiness and excitement. These

emotions are typically associated with joy, contentment, and enthusiasm, and are often characterized by a sense of well-being and positivity.

The aim of testing Wysa, the chatbot, was to evaluate its ability to accurately detect the user's emotions based on this classification spanning tree. By identifying and categorizing emotions into major and minor cases, Wysa aims to provide insights and support to users by understanding their emotional state and providing appropriate responses and guidance. Through this evaluation, it can be determined if Wysa is effective in accurately detecting and responding to the user's emotions, thus offering valuable insights into the capabilities of the chatbot in the field of emotion detection and support.



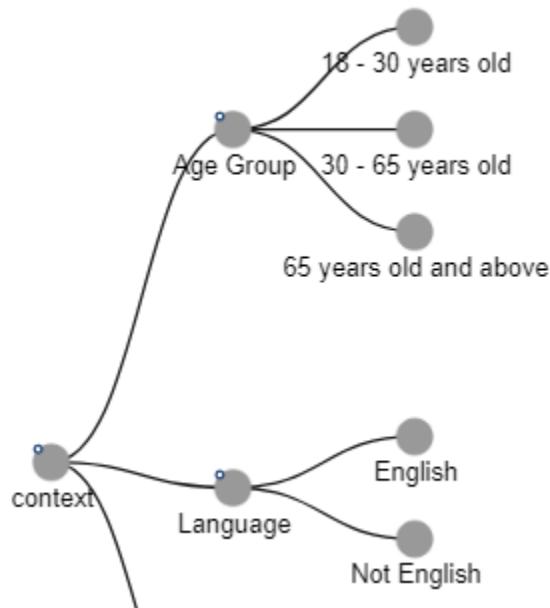
## Context

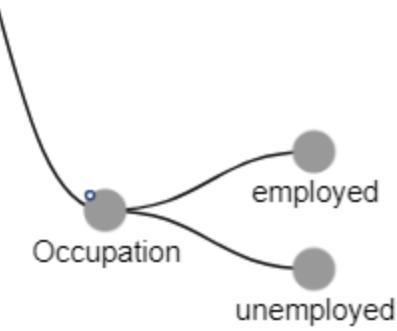
The context classification spanning tree for feeling detection encompasses four major cases and ten minor cases, providing a comprehensive framework to evaluate Wysa's performance in various contextual scenarios. The first major case focuses on age groups, encompassing three distinct age ranges: 18 to 30 years old, 30 to 65 years old, and 65+ years old. This allows for testing Wysa's functionality and effectiveness across different age groups, considering potential variations in emotional expression and needs across different age groups.

The second major case in the spanning tree is language, aiming to assess Wysa's ability to understand languages other than English. This case recognizes the importance of linguistic diversity and evaluates Wysa's performance in comprehending and responding to emotions expressed in languages other than English..

The last major case in the spanning tree is occupation, encompassing the user's occupation as a contextual factor. This case recognizes that an individual's occupation could impact their emotions, and evaluates Wysa's ability to understand and respond to emotions expressed in different professional contexts.

The seven minor cases within the major cases of age group, language, and occupation further provide an evaluation of Wysa's performance in various specific scenarios within each major case. This structured approach to testing allows for a comprehensive assessment of Wysa's capabilities in different contexts, providing insights into its effectiveness and potential areas for improvement in detecting and addressing emotions across diverse populations.

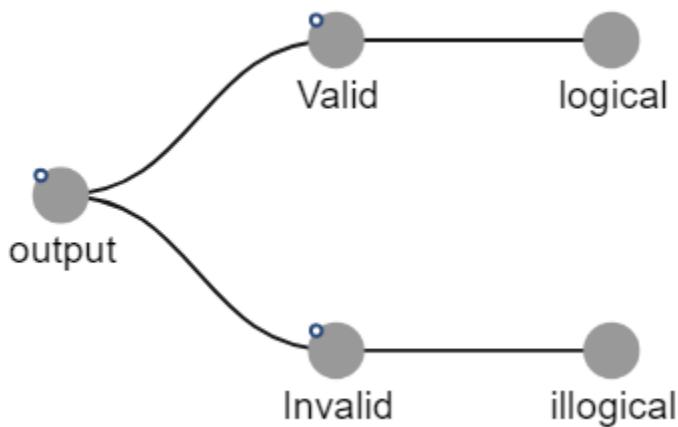




## Output

The output classification spanning tree for feeling detection comprises two major cases and two minor cases, providing a framework for evaluating the accuracy and logic of its responses. The first major case is labeled as "valid," which encompasses responses that are logical and coherent, aligning with the expected output from Wysa. These responses are deemed accurate and reliable, indicating that Wysa is providing meaningful and relevant answers to user queries.

The second major case in the classification tree is labeled as "invalid," encompassing responses that are illogical. The minor case of "illogical" includes responses that may not make sense or lack coherence, indicating that Wysa may not be providing logical answers.



Context model decision table

| ID | Age Group         | Language    | Occupation |
|----|-------------------|-------------|------------|
| C1 | 18 - 30 years old | -           | -          |
| C2 | -                 | Not English | -          |
| C3 | -                 | -           | Employed   |

Input model Decision table

| ID | Good Feelings | Bad Feelings |
|----|---------------|--------------|
| I1 | Happy         | -            |
| I2 | Excited       | -            |
| I3 | -             | Stress       |
| I4 | -             | Sad          |
| I5 | -             | Anxious      |

Output model decision table

| ID | Output    |
|----|-----------|
| O1 | Logical   |
| O2 | Illogical |

Test case design table

I have devised 8 test cases that encompass both the input and context model spanning trees. Each minor case will have a corresponding test case to ensure that Wysa accurately detects feelings, whether they are positive or negative, in various contextual scenarios. The first two test cases will focus on feeling detection with both invalid and valid conversation formats. The subsequent three test cases will solely test the accuracy of feeling detection. Lastly, the last three test cases will evaluate both

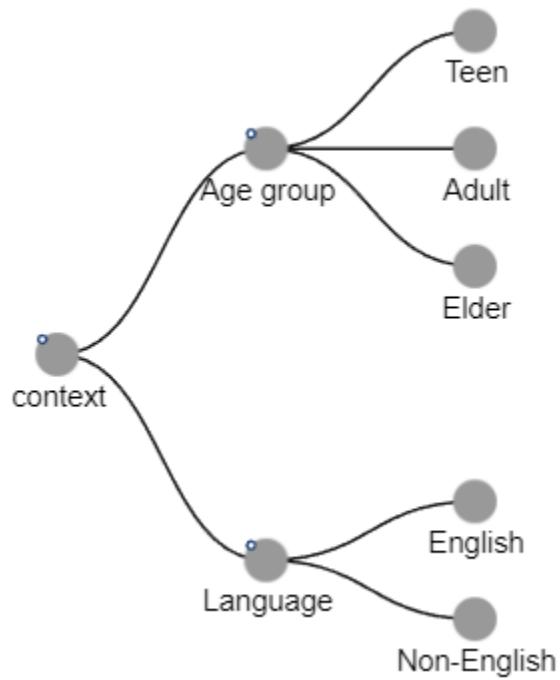
feeling detection and contextual scenarios, such as age group, language, and occupation.

| Test Case ID | Input | Context | Output |
|--------------|-------|---------|--------|
| 1            | I1    | -       | O1     |
| 2            | I2    | -       | O2     |
| 3            | I3    | -       | O1     |
| 4            | I4    | -       | O1     |
| 5            | I5    | -       | O1     |
| 6            | I1    | C1      | O2     |
| 7            | I3    | C2      | O2     |
| 8            | I5    | C3      | O2     |

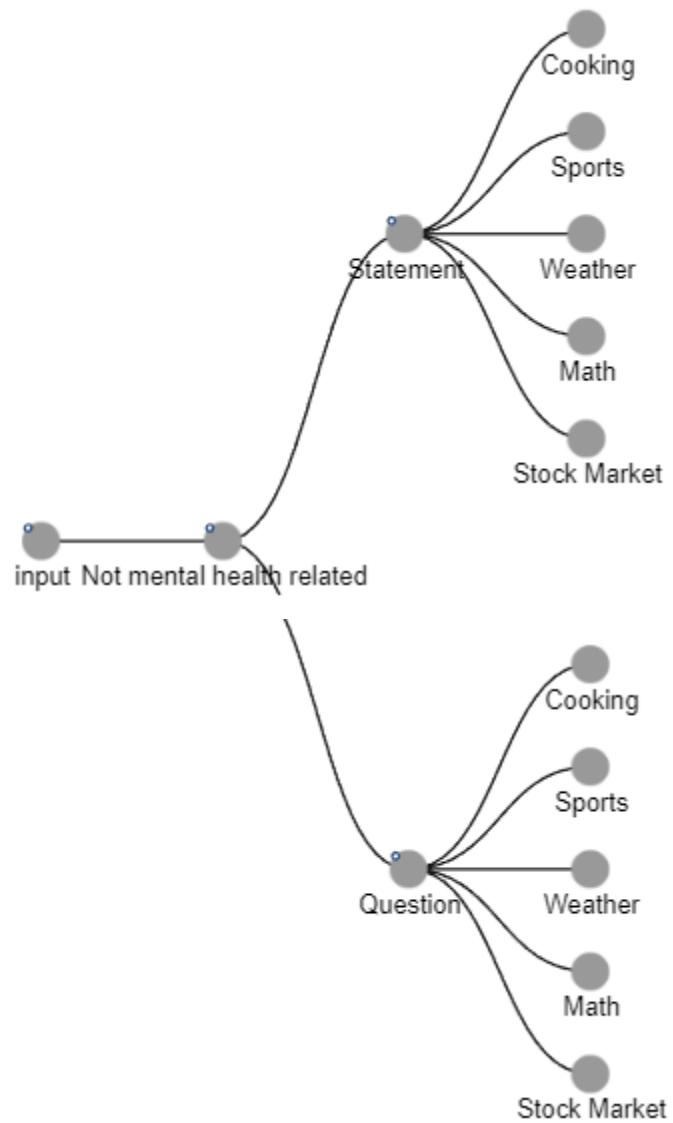
## Invalid Topics - Kevin

In order to test Wysa's AI capacity we also opted to test how it would handle topics unrelated to mental health.

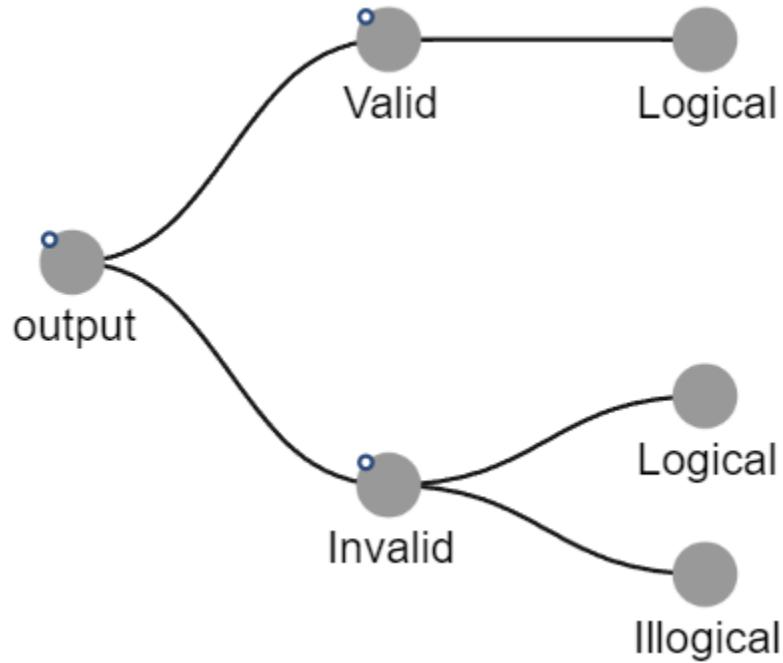
## Ai-Powered Function Context Classification



## Ai-Powered Function Input Classification



## Ai-Powered Function Output Classification



## Ai-Powered Function Classification Decision Table

[Input Decision Table](#)

| @id | Not mental health related.Question | Not mental health related.Statement |
|-----|------------------------------------|-------------------------------------|
| 1   | Cooking                            | Cooking                             |
| 2   | Stock Market                       | Sports                              |
| 3   | Cooking                            | Weather                             |
| 4   | Sports                             | Weather                             |
| 5   | Weather                            | Weather                             |
| 6   | Math                               | Weather                             |
| 7   | Stock Market                       | Weather                             |
| 8   | Cooking                            | Math                                |
| 9   | Sports                             | Math                                |
| 10  | Weather                            | Math                                |
| 11  | Math                               | Math                                |
| 12  | Sports                             | Cooking                             |
| 13  | Stock Market                       | Math                                |
| 14  | Cooking                            | Stock Market                        |
| 15  | Sports                             | Stock Market                        |
| 16  | Weather                            | Stock Market                        |
| 17  | Math                               | Stock Market                        |
| 18  | Stock Market                       | Stock Market                        |
| 19  | Weather                            | Cooking                             |
| 20  | Math                               | Cooking                             |
| 21  | Stock Market                       | Cooking                             |
| 22  | Cooking                            | Sports                              |
| 23  | Sports                             | Sports                              |
| 24  | Weather                            | Sports                              |
| 25  | Math                               | Sports                              |

### Context Decision Table

| @id | Age group | Language    |
|-----|-----------|-------------|
| 1   | Teen      | English     |
| 2   | Teen      | Non-English |
| 3   | Adult     | English     |
| 4   | Adult     | Non-English |
| 5   | Elder     | English     |
| 6   | Elder     | Non-English |

Output Decision Table

| @id | Invalid   | Valid     |
|-----|-----------|-----------|
| 1   | undefined | undefined |
| 2   | undefined | undefined |
| 3   | undefined | undefined |
| 4   | undefined | undefined |
| 5   | undefined | undefined |
| 6   | undefined | undefined |
| 7   | undefined | undefined |
| 8   | undefined | undefined |
| 9   | undefined | undefined |
| 10  | undefined | undefined |
| 11  | undefined | undefined |
| 12  | undefined | undefined |
| 13  | undefined | undefined |
| 14  | undefined | undefined |
| 15  | undefined | undefined |
| 16  | undefined | undefined |
| 17  | undefined | undefined |
| 18  | undefined | undefined |
| 19  | undefined | undefined |
| 20  | undefined | undefined |
| 21  | undefined | undefined |
| 22  | undefined | undefined |
| 23  | undefined | undefined |
| 24  | undefined | undefined |
| 25  | undefined | undefined |

# Section 3 – AI Function Test Cases with Inputs/Expected Outputs

AI-powered function test sets

In addition to doing the test cases, we also use the AI testing tool to augment the test case inputs, that is modified the inputs to test whether or not Wysa can still detect the meaning of those sentences when the input is modified in different ways, like missing words, random characters, etc.

## Mental Issues Function

### Test case #1

|                        |  |  |
|------------------------|--|--|
| Test Case ID           | M1   |  |
| Test Case # in AI Tool | 37   |  |
| Test Case Description  | Tests if Wysa can understand the User's mental health and thoughts to suggest ways to help the User feel better. |  |
| Input                  | Mental Issue: Depression, Cause: Thought   |  |
| Steps #                | Wysa   | User   |
| 1                      | Asks User how they are feeling   |  |
| 2                      |  | User says "I'm feeling depressed"              |
| 3                      | Asks User what feelings made me feel depressed   |  |
| 4                      |  | User says "Bad thoughts"                       |
| 5                      | Asks User to share one thought about the situation   |  |
| 6                      |  | User states a thought                          |
| 7                      | Lists thoughts that might cause the way the User is feeling  |  |
| 8                      |  | User chooses which thought they relate to most |
| 9                      | Guides User on ways to make them   |  |

|                 |  |                    |
|-----------------|--|--------------------|
|                 | feel better based on the User's thought                                  |                    |
| 10              |  | User follows guide |
| Expected Output | Wysa recommends a guide for the User to follow to make them feel better. |                    |

Augmentation test cases

Test Case 1.1

Step 2: random word swap: “depressed feeling I’m”

Test Case 1.2

Step 2: random word delete: “feeling depressed”

## Test case #2

Short term memory about User's stress and thoughts

|                        |   |  |
|------------------------|---|--|
| Test Case ID           | M2  |  |
| Test Case # in AI Tool | 46  |  |
| Test Case Description  | Tests the short term memory of Wysa by testing if it remembers the User's mental health information and thoughts. |  |
| Input                  | Mental Issue: Stress, Cause: Thought  |  |
| Steps #                | Wysa  | User   |
| 1                      | Asks User how they are feeling  |  |
| 2                      |   | Says “I'm feeling stressed”                                    |
| 3                      | Asks if User is anxious   |  |
| 4                      |   | User responds “Yes”  |
| 5                      | Asks if it is affecting the User physically too   |  |
| 6                      |   | User can choose “Mostly mind” Mostly physical” “A bit of both” |
| 7                      |   | User chooses “Mostly mind”                                     |

|                  |  |  |
|------------------|--|--|
| 8                | Asks User what thought comes to mind that causes their stress  |  |
| 9                |  | User says "I have a lot of homework"     |
| 10               | Asks the user what change they would hope for  |  |
| 11               |  | User says what they would like to change |
| 12               | Restates the user's thought, "I have a lot of homework" and asks what positives about the situation/themselves have been ignored |  |
| 13               |  | User creates positive thought            |
| Expected Outcome | Wysa should be able to remember User's mental health and thought input to help User feel better.                                 |  |

### Test case #3

|                         |   |                             |
|-------------------------|---|-----------------------------|
| Test Case ID            | M3  |                             |
| Test # on AI in AI Tool | 47  |                             |
| Test Case Description   | Tests the long term memory of Wysa by testing if it remembers the User's mental health information and event. |                             |
| Input                   | Mental issue: Stress, Cause: Event  |                             |
| Steps #                 | Wysa  | User                        |
| 1                       | Asks User how they are feeling  |                             |
| 2                       |   | Says "I'm feeling stressed" |
| 3                       | Asks if User is anxious   |                             |
| 4                       |   | User responds "Yes"         |
| 5                       | Asks if it is affecting the User physically too   |                             |

|                  |  |  |
|------------------|--|--|
| 6                |  | User can choose “Mostly mind” Mostly physical” “A bit of both” |
| 7                |  | User chooses “Mostly mind”                                     |
| 8                | Asks User what event causes their stress   |  |
| 9                |  | User says “I have a midterm today”                             |
| 10               | Asks the user what change they would hope for  |  |
| 11               |  | User says what they would like to change                       |
| 12               | Restates the user’s thought, “I have a lot of homework” and asks what positives about the situation/themselves have been ignored |  |
| 13               |  | User creates positive thought                                  |
| 14               | Asks User if they feel better  |  |
| 15               |  | User says “Yes”  |
| 16               | Prompts the User to “Start new Chat”   |  |
| 17               |  | User presses “Start new Chat”                                  |
| 18               | Wysa states that User has mentioned that they’re “feeling stressed”  |  |
| Expected Outcome | Wysa should be able to remember the User’s feeling of stress and event that causes it even after the User creates a new chat     |  |

#### Test case #4

|                        |  |
|------------------------|--|
| Test Case ID           | M4   |
| Test Case # in AI Tool | 104  |
| Test Case Description  | Tests Wysa’s ability to respond to User’s conversation format that contains consecutive questions. |

|                  |   |   |
|------------------|---|---|
| Input            | Mental Issue: Sleep, Cause: Event   |   |
| Steps #          | Wysa  | User  |
| 1                |   | Asks "How can I fix my sleep schedule?"     |
| 2                | Offers ways to help manage User's sleep schedule  |   |
| 3                |   | Asks "When was the last time I slept?"      |
| 4                | Tells User the time which they last slept   |   |
| 5                |   | Asks "Why am I not sleeping well at night?" |
| 6                | Tells User why they can't sleep   |   |
| 7                | Offers User good sleeping habits and recommendations  |   |
| Expected Outcome | Wysa should be able to answer all the User's questions and offers recommendations to fix User's sleeping habits |   |

### Test case #5

|                        |  |  |
|------------------------|--|--|
| Test Case ID           | M5   |  |
| Test Case # in AI Tool | 101  |  |
| Test Case Description  | Tests Wysa's ability to recognize change in conversation |  |
| Input                  | Mental Issue: Sleep, Cause: Feeling                      |  |
| Steps #                | Wysa   | User   |
| 1                      | Asks how is the User's day                               |  |
| 2                      |  | User responds "I am feeling tired."            |
| 3                      | Instructs user to tell Wysa about their day              |  |
| 4                      |  | User asks, "Why am I having trouble sleeping?" |

|                  |  |   |
|------------------|--|---|
| 5                | User provides a list of reasons of why the User might be having trouble sleeping           |   |
| 6                |  | User chooses which reason in the list applies to them |
| 7                | Wysa recommends ways to help them sleep better   |   |
| Expected Outcome | Wysa should be able to adjust to the User's conversation format and answer their question. |   |

### Test case #6

|                        |   |                                     |
|------------------------|---|-------------------------------------|
| Test Case ID           | M6  |                                     |
| Test Case # in AI Tool | 15  |                                     |
| Test Case Description  | Tests Wysa's ability to recognize the User's instruction after asking a question.                       |                                     |
| Input                  | Mental Issue: Depression, Cause: Feeling  |                                     |
| Steps #                | Wysa  | User                                |
| 1                      | Asks User how they are feeling today  |                                     |
| 2                      |   | User responds "I am feeling tired." |
| 3                      | Asks User if there is a thought that makes them feel this way   |                                     |
| 4                      |   | User says "I am depressed"          |
| 5                      | Asks if they can offer recommendations to help the User's depression                                    |                                     |
| 6                      | Offers recommendations that will help the User's mood   |                                     |
| 7                      |   | User follows recommendations        |
| Expected Outcome       | Wysa should be able to adjust to the User's conversation format and respond to the User's mental issue. |                                     |

## Test case #7

|                        |   |   |
|------------------------|---|---|
| Test Case ID           | M7  |   |
| Test Case # in AI Tool | 16  |   |
| Test Case Description  | Tests Wysa's ability to end a series of questions to find out what causes the User's mental issue       |   |
| Input                  | Mental Issue: Anxiety, Cause: Feelings  |   |
| Steps #                | Wysa  | User  |
| 1                      | Asks how the User is feeling  |   |
| 2                      |   | User says, "What are different exercises that will help with my anxiety?" |
| 3                      | Asks what the cause of their anxiety  |   |
| 4                      |   | User states that negative feelings is the cause of their anxiety          |
| 5                      | Provides a list of types of feelings that the user might be going through                               |   |
| 6                      |   | User chooses the feeling that they relate to most                         |
| 7                      | Asks to confirm if this feeling is what they are feeling  |   |
| 8                      |   | User says, "Let's do something else"                                      |
| 9                      | Asks User what they want to do  |   |
| Expected Outcome       | Wysa should be able to adjust to the User's conversation format and respond to the User's mental issue. |   |

## Test case #8

|                        |    |
|------------------------|----|
| Test Case ID           | M8 |
| Test Case # in AI Tool | 16 |

|                       |  |                                     |
|-----------------------|--|-------------------------------------|
| Test Case Description | Tests if Wysa can understand the User in a different language that is not english.             |                                     |
| Input                 | Mental Issue: Anxiety, Cause: Event  |                                     |
| Steps #               | Wysa   | User                                |
| 1                     | Asks how the User is feeling   |                                     |
| 2                     |  | Says, “私は不安を感じています”                 |
| 3                     | Asks what the cause of their anxiety   |                                     |
| 4                     |  | Says, “トラウマ的な出来事”                   |
| 5                     | Asks what event causes their anxiety   |                                     |
| 6                     |  | Says “友達が死んだ”                       |
| 7                     | Says they understand and offers ways to help the User's anxiety                                |                                     |
| 8                     |  | User follows Wysa's recommendations |
| Expected Outcome      | Wysa should be able to understand the User even when they are speaking in a different language |                                     |

### Test case #9

|                        |  |                     |
|------------------------|--|---------------------|
| Test Case ID           | M9   |                     |
| Test Case # in AI Tool | 41   |                     |
| Test Case Description  | Tests if Wysa can understand the User in a different language that is not english. |                     |
| Input                  | Mental Issue: Anxiety, Cause: Event  |                     |
| Steps #                | Wysa   | User                |
| 1                      | Asks how the User is feeling   |                     |
| 2                      |  | Says, “私は不安を感じています” |
| 3                      | Asks what the cause of their anxiety   |                     |

|                  |  |                                     |
|------------------|--|-------------------------------------|
| 4                |  | Says, “トラウマ的な出来事”                   |
| 5                | Asks what event causes their anxiety   |                                     |
| 6                |  | Says “友達が死んだ”                       |
| 7                | Says they understand and offers ways to help the User's anxiety                                |                                     |
| 8                |  | User follows Wysa's recommendations |
| Expected Outcome | Wysa should be able to understand the User even when they are speaking in a different language |                                     |

### Test case #10

|                        |   |                                     |
|------------------------|---|-------------------------------------|
| Test Case ID           | M10   |                                     |
| Test Case # in AI Tool | 100   |                                     |
| Test Case Description  | Tests if Wysa can provide recommendations that will help the User's Anxiety based on their thoughts |                                     |
| Input                  | Mental Issue: Anxiety, Cause: Thought   |                                     |
| Steps #                | Wysa  | User                                |
| 1                      | Asks how the User is feeling  |                                     |
| 2                      |   | Says they have anxiety              |
| 3                      | Asks what the cause of their anxiety  |                                     |
| 4                      |   | Says, “Bad thoughts”                |
| 5                      | Asks what thoughts causes their anxiety   |                                     |
| 6                      |   | Lists their thoughts                |
| 7                      | Says they understand and offers ways to help the User's anxiety                                     |                                     |
| 8                      |   | User follows Wysa's recommendations |

|                  |   |
|------------------|---|
| Expected Outcome | Wysa should provide recommendations based on the User's thoughts. |
|------------------|---|

### Test case #11

|                        |   |  |
|------------------------|---|--|
| Test Case ID           | M11   |  |
| Test Case # in AI Tool | 14  |  |
| Test Case Description  | Tests if Wysa can provide recommendations to a Teenager and is Employed who is suffering from depression. |  |
| Input                  | Mental Issue: Depression, Cause: Events   |  |
| Steps #                | Wysa  | User   |
| 1                      | Asks how the User is feeling  |  |
| 2                      |   | Says, "I am depressed"                         |
| 3                      | Asks what the cause of their depression   |  |
| 4                      |   | Says, "I am being bullied at school"           |
| 5                      | Asks what other thoughts or events they are thinking or going through.                                    |  |
| 6                      |   | Lists events of reasons why they are depressed |
| 7                      | Says they understand and offers ways to help the User's depression  |  |
| 9                      |   | User follows Wysa's recommendations            |
| Expected Outcome       | Wysa should provide recommendations based on the User's context and causes of events.                     |  |

### Test case #12

|                        |     |
|------------------------|-----|
| Test Case ID           | M12 |
| Test Case # in AI Tool | 14  |

|                       |   |   |
|-----------------------|---|---|
| Test Case Description | Tests if Wysa can provide recommendations to an Elder and is Unemployed who is suffering from depression.                     |   |
| Input                 | Mental Issue: Sleep, Cause: Thoughts  |   |
| Steps #               | Wysa  | User  |
| 1                     | Asks how the User is feeling  |   |
| 2                     |   | Says, "I am lacking in sleep"                         |
| 3                     | Asks what thoughts they have that causes their sleeping problem   |   |
| 4                     |   | Says, "I am overthinking too much"                    |
| 5                     | Says they understand and offers exercises and relaxing recommendation that can help them sleep                                |   |
| 6                     |   | User follows recommendations                          |
| 7                     | Suggests a reminder to set at night to help remind the User that it is time to sleep  |   |
| 9                     |   | User sets the sleeping time that will best suit them. |
| Expected Outcome      | Wysa should be able to provide reminders that will help remind the User to sleep at a certain time and other recommendations. |   |

## Exercise function

The detailed test cases are outlined as below. The User column shows what the user actions and text messages. The Wysa column shows what is the expected output of Wysa.

### Test case #1

Normal happy test case, this is the base case to demonstrate how Wysa works

| Steps # | Wysa                                   | User                              |
|---------|--|-----------------------------------|
| 1       |  | “Let’s get active”                |
| 2       | Wysa offers some exercise choices      |                                   |
| 3       |  | User chooses one of the exercises |
| 4       | Wysa explains the exercises in details |                                   |
| 5       |  | User follows the exercises        |
| 6       |  | “Let’s get mindfulness”           |
| 7       | Wysa offers mindfulness exercises      |                                   |
| 8       |  | User chooses one of the exercises |
| 9       | Wysa explains the exercises in details |                                   |
| 10      |  | User follows the exercises        |
| 11      | Wysa asks for feedback                 |                                   |
| 12      |  | User gives some kind of feedback  |

Expected output: Wysa should pass and recognize user's response

## Test case #2

This test case will built upon the base case with context input to test whether or not Wysa offer workout personalization for different demographic group

Test case 237 in AI testing tool

| Steps # | Wysa | User                    | Note |
|---------|------|-------------------------|------|
| 1       |      | “Give me some exercises |      |

|    |  |  |                        |
|----|--|--|------------------------|
|    |  | for elders”                                  |                        |
| 2  | Wysa offers some physical exercises choices          |  |                        |
| 3  |  | User chooses one of the exercises            |                        |
| 4  | Wysa explains the exercises in details               |  |                        |
| 5  |  | User follows the exercises                   |                        |
| 6  |  | “Give me some exercises for kids”            |                        |
| 7  | Wysa offers some physical exercises choices for kids |  |                        |
| 8  |  | User chooses one of the exercises            |                        |
| 9  | Wysa explains the exercises in details               |  |                        |
| 10 |  | User follows the exercises                   |                        |
| 11 |  | User asks the app what exercises did they do | Test memory of the app |
| 12 | Wysa responds with correct answer                    |  |                        |

Augmentation test cases

Input at step 1 is test case #237 in the testing tool

Input at step 6 is test case #15 in the testing tool

Test case 2.1:

Step 1: random word swap: "I to want do some exercises as an elder"

Step 6: random word delete: "Give me exercises appropriate for Asian kids"

Test case 2.2

Step 1: keyboard augmentation: "I wanr to do some exerciX3s as an elder"

Step 6: random word insert: "Give me exercises appropriate example for Asian kids"

Wysa expected to still able to understand the main meaning behind or text or ask users to clarify what they just said

Test case #3

Keep changing exercise choices and see whether or not Wysa can handle it

| Steps # | Wysa  | User   | Note |
|---------|---|--|------|
| 1       |   | "Give me some exercises for mindfulness and workout for Latino kids" |      |
| 2       | Wysa asks user to clarify which exercises they want to do |  |      |
| 3       |   | "Give me some mindfulness exercises for Latino adult"                |      |
| 4       | Wysa offers some tailored mindfulness exercises           |  |      |
| 5       |   | "Give me some physical"  |      |

|   |   |   |                                     |
|---|---|---|-------------------------------------|
|   |   | exercises for Latino adult”                               |                                     |
| 6 | Wysa offers some tailored physical exercises  |   |                                     |
| 7 |   | “I want to express my thoughts about what happened today” | Test conversation format of the app |
| 8 | Wysa should either:<br>1. Ask to clarify what did the user asks<br>2. Change topic and ask user to clarify their thoughts |   |                                     |

Augmentation test cases

Input at step 1 is test case 89

Input at step 3 is test case 121

Input at step 5 is test case 125

Test case 3.1:

Step 1: random char swap: “Giev me some execirses for mindfulness and workout for Laitno kids”

Step 3: random word swap: “Give some me mindfulness exercises for Latino adult”

Step 5: synonym augmentation: “Establish me some physical exercises for Latino adult”

Test case 3.2:

Step 1: random char swap: “Give me soe exercises for mndflnss and workout for Latno kids”

Step 3: ocr augmentation: “Cive me some mindfulness exerci8e8 for Latino adult”

Step 5: ocr augmentation: “Cive me some physical exercises for Latino adu1t”

Test case 3.3:

Step 1: ocr augmentation: "Give me some exercises for mindfulness and workout f0r Latino kids"

Step 3: random char insert: "Giv3e me some mindfulness dexercisies for Latino adult"

Step 5: keyboard augmentation: "Gove me some physical exedcisFs for Latino adult"

## Test case #4

Test case with different language

- User asks for workout in a different language

If Wysa can support other language, it should respond in other language

| Steps # | Wysa  | User  |
|---------|---|---|
| 1       |   | "Tôi muốn tập thể dục" (I want to exercise but in Vietnamese) |
| 2       | Wysa should either:<br>1. Respond in the appropriate language<br>2. Tell the user to respond in English as it currently does not support this language or other language than English |   |

## Test case #5:

Ask Wysa to set exercise frequency

| Steps # | Wysa | User | Note |
|---------|------|------|------|
|         |      |      |      |

|   |   |   |  |
|---|---|---|--|
| 1 |   | "I want to set reminder to do some mental exercises everyday" |  |
| 2 | Wysa responds with prompts to guide user to set up exercise frequency |   |  |
| 3 |   | User follows the prompt                                       |  |
| 4 | Wysa set up exercises reminder  |   |  |

Input at step 1 is test case 5

Augmented test case

Test case 5.1:

Step 1: synonym augmentation: I want to set reminder to coiffe some mental exercises everyday

## Feeling detection

|                 |  |
|-----------------|--|
| Test case ID    | 1  |
| Description     | Happiness detection with valid conversation format   |
| Purpose         | To test Wysa to see if Wysa could detect happiness   |
| Input           | "I'm feeling so happy today!"  |
| Output          | "I'd love to hear about what makes you feel happy today"<br>"What would make today better" |
| Expected Output | "That's great to hear! What made you happy today?"   |
| Result          | Passed   |

|              |   |
|--------------|---|
| Test case ID | 2 |
|--------------|---|

|                 |  |
|-----------------|--|
| Description     | excitement detection with invalid conversation format  |
| Purpose         | To test Wysa to see if Wysa could detect excitement  |
| Input           | "I won the lottery today!"<br>"I just got promoted at work!"   |
| Output          | "What would you require to focus on that?"<br>""   |
| Expected Output | "I'm having trouble understanding your message. Can you please provide more details or follow a proper conversation format?" |
| Result          | failed   |

|                 |   |
|-----------------|---|
| Test case ID    | 3   |
| Description     | Stress detection  |
| Purpose         | To test Wysa to see if Wysa could detect stress   |
| Input           | "Work has been really overwhelming lately, and I'm feeling stressed."   |
| Output          | "That's an uncomfortable feeling, I know. Observe how it feels for a bit, just allowing it to exist."                                     |
| Expected Output | "It sounds like you're going through a stressful time with your exams. Remember to take breaks and practice self-care during this period" |
| Result          | Passed  |

|              |   |
|--------------|---|
| Test case ID | 4   |
| Description  | Sad detection   |
| Purpose      | To test Wysa to see if Wysa could detect sadness                      |
| Input        | "I'm feeling really sad today"  |
| Output       | "I understand you're feeling sad. It's only natural to feel this way" |

|                 |  |
|-----------------|--|
| Expected Output | "I'm sorry to hear that you're feeling sad. Remember to take care of yourself and reach out to supportive people around you" |
| Result          | Passed   |

|                 |  |
|-----------------|--|
| Test case ID    | 5  |
| Description     | Anxiety detection  |
| Purpose         | To test Wysa to see if Wysa could detect anxiety   |
| Input           | "I'm feeling really anxious about my presentation tomorrow"  |
| Output          | "I'm here to help you feel more in control"  |
| Expected Output | "It sounds like you're feeling anxious about your presentation. Remember to practice deep breathing and positive self-talk to manage your anxiety" |
| Result          | Passed   |

|                 |   |
|-----------------|---|
| Test case ID    | 6   |
| Description     | Happy detection with context of a person who is from 18 - 30 years old  |
| Purpose         | To test Wysa to see if Wysa could detect happiness expressed by a user who is from 18 to 30 years old   |
| Input           | "I aced my exam and I'm so happy!"  |
| Output          | "Scientists say that positive emotions increase one's awareness and encourage novel thoughts and actions. This could be a fantastic day to start something creative." |
| Expected Output | "Congratulations on acing your exam! That's awesome! How do you plan to celebrate?"   |
| Result          | Failed  |

|                 |   |
|-----------------|---|
| Test case ID    | 7   |
| Description     | Stress detection in non English language  |
| Purpose         | To test Wysa to see if Wysa could detect stress in non English language   |
| Input           | "Estoy muy estresado por los exámenes" (Spanish for "I am very stressed about exams").  |
| Output          | "That looks like a bunch of letters that don't make a lot of sense together"<br>"Are you testing my AI capabilities?"   |
| Expected Output | "Comprendo que te sientas estresado(a) por los exámenes. Trata de descansar y respirar profundamente" (Spanish for "I understand that you feel stressed about exams. Try to rest and breathe deeply") |
| Result          | Failed  |

|                 |   |
|-----------------|---|
| Test case ID    | 8   |
| Description     | Anxiety for a employed user   |
| Purpose         | To test Wysa to see if Wysa could detect anxiety expressed by a employed user   |
| Input           | "I'm feeling really anxious about the upcoming presentation at work"  |
| Output          | "This kind of distress can feel overwhelming. I can help you learn ways to feel more in control"                            |
| Expected Output | "It's understandable to feel anxious about the upcoming presentation. Remember to take deep breaths and practice self-care" |
| Result          | Failed  |

## Invalid Input

### Test case #1

| Steps # | Wysa   | User               |
|---------|--|--------------------|
| 1       | Let's talk about your day.                               |                    |
| 2       | What made you feel okay?                                 |                    |
| 3       |  | What is 8 squared? |
| 4       | Let's talk about that. Go on...                          |                    |
| 5       |  | What is 2+2?       |
| 6       | I'm not great with questions, but I can listen.. Go on.. |                    |

### Test case #2

| Steps # | Wysa  | User                                       |
|---------|---|--|
| 1       | What would you like to do now?  |  |
| 2       |   | How are GameStop stocks looking right now? |
| 3       | Research suggests a link between positive emotions and success.<br>What a great day to achieve something. |  |
| 4       | What's the main focus for today?  |  |

### Test case #3

| Steps # | Wysa                           | User |
|---------|--------------------------------|------|
| 1       | What would you like to do now? |      |

|   |   |   |
|---|---|---|
| 2 |   | What temperature is it safe to eat chicken? |
| 3 | I function strictly within the boundaries set by our privacy practices. All our conversations are private. You can read more about our privacy practices on our website: <a href="http://www.wysa.io">www.wysa.io</a> |   |
| 4 | Do you have anything else to ask me?  |   |

#### Test case #4

| Steps # | Wysa   | User                         |
|---------|--|------------------------------|
| 1       | What would you like to do now?   |                              |
| 2       |  | Who are the New York Giants? |
| 3       | Research suggests a link between positive emotions and success. What a great day to achieve something. |                              |
| 4       | What's the main focus for today?   |                              |

#### Test case #5

| Steps # | Wysa  | User                   |
|---------|---|------------------------|
| 1       | What would you like to do now?  |                        |
| 2       |   | Will it rain Thursday? |
| 3       | Scientists say that positive emotions increase one's awareness and encourage novel thoughts and actions. This could be a fantastic day to start |                        |

|   |                                  |  |
|---|----------------------------------|--|
|   | something creative.              |  |
| 4 | What's the main focus for today? |  |

## Section 4 - Test complexity

### Feeling function

Test case number:

- There are 12 test cases, 2 augmented test cases
- The test complexity is 14
- Each test covers the following topics:
  - Input: every type of input at least 1 test case covered
  - Context:
    - Age group: Teen, Adult, Elder
    - Occupation: Student, Employee, Unemployed

### Exercise function

Test case number:

- There are 5 test cases, 6 augmented test cases, therefore test complexity is 11
- There are at least 1 test covering these topics:
  - Input: every type of input at least 1 test case covered
  - Context:
    - Age group: Kid, adult, elder
    - Demographics: Asian, Latino

**Deliverable 3 - AI Test Automation Report**

**CMPE 187 - Software Quality Engineering**

**San Jose State University**

**Group 2**

Team members:

Khai Nguyen

Faith Deanon

Kevin Campbell

Long Nguyen

# 1. AI Testing Tool demo

## Data augmentation

|           |           |           |           |                         |
|-----------|-----------|-----------|-----------|-------------------------|
| undefined | undefined | undefined | undefined | <button>Upload</button> |
| undefined | undefined | undefined | undefined | <button>Upload</button> |

First click on Upload button for each test case

### Upload Text

Please enter your input text:

Give me some mindfulness exercises for Asian elders

Save ✓ Cancel ✕

Then enter your text input

|                                     |   |                       |      |                  |          |      |
|-------------------------------------|---|-----------------------|------|------------------|----------|------|
| <input checked="" type="checkbox"/> | 3 | Set exercise schedule | Good | Reframe thoughts | Strength | Teen |
|-------------------------------------|---|-----------------------|------|------------------|----------|------|

Then check the box on the test case

|  |   |   |   |   |   |   |  |    |   |   |
|--|---|---|---|---|---|---|--|----|---|---|
| «  | < | 1 | 2 | 3 | 4   | 5 | ...  | 20 | > | » |
| Select rows per page ▾                     |   |   |   |   |   |   |  |    |   |   |
| <input type="button" value="Import Data"/> |   |   |   |   | <input type="button" value="Export Testcases"/> |   | <input type="button" value="Data Augmentation"/> |    |   |   |

At below click on Data augmentation

## Data Augmentation

### Text test cases

| Test Case ID | Input Text  |
|--------------|---|
| 3            | Give me some mindfulness exercises for Asian elders |

### Augment Options

Select All

- random word swap
- random word delete
- random word insert
- synonym augmentation
- ocr augmentation
- keyboard augmentation
- random char insert
- random char swap
- random char delete

[Add test case](#)



[Back to Testcase Table](#)

## Choose any data augmentation options

### Data Augmentation

#### Text test cases

| Test Case ID | Input Text  |
|--------------|---|
| 3            | Give me some mindfulness exercises for Asian elders |

#### Augment Options

Select All

- random word swap
- random word delete
- random word insert
- synonym augmentation
- ocr augmentation
- keyboard augmentation
- random char insert
- random char swap
- random char delete

#### Augmented Test Cases

| Base Test Case ID | Input Text  | Augment Profile    | Augment Result  |
|-------------------|---|--------------------|---|
| 3                 | Give me some mindfulness exercises for Asian elders | random word swap   | Give some me mindfulness exercises for Asian elders             |
| 3                 | Give me some mindfulness exercises for Asian elders | random word insert | Give heedfulness me some mindfulness exercises for Asian elders |

[Add test case](#)



[Back to Testcase Table](#)

The tool will generate the augmented test case  
Click on Back to Testcase Table to go back

|                                     |   |                       |      |                  |          |      |
|-------------------------------------|---|-----------------------|------|------------------|----------|------|
| <input checked="" type="checkbox"/> | 3 | Set exercise schedule | Good | Reframe thoughts | Strength | Teen |
|-------------------------------------|---|-----------------------|------|------------------|----------|------|

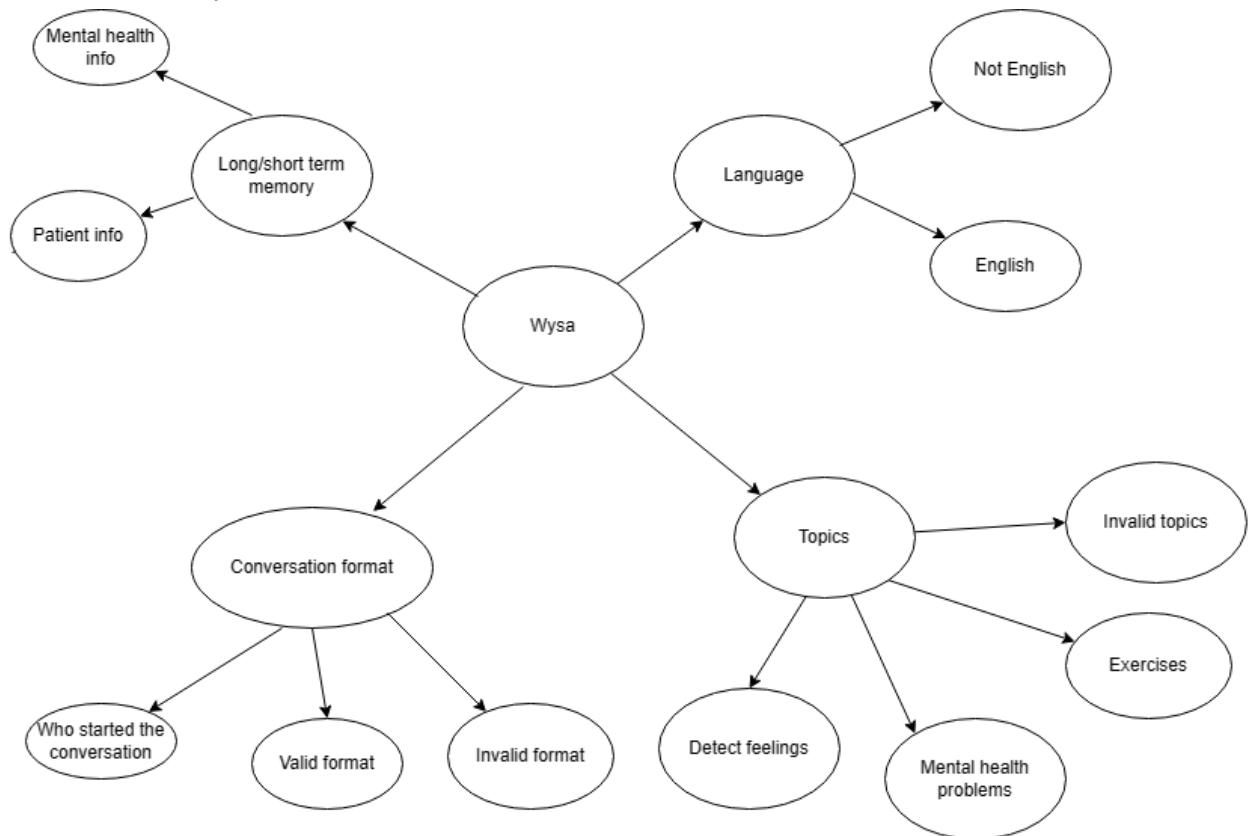
id: 1 augment method: random word swap

id: 2 augment method: random word insert

Everytime you check the text case, the augmented test case will be shown

## 2. Test scripts

We tested the Wysa app based on 4 topics:



- Those are Invalid topics, exercises topic, mental health problems and detect feelings topic
- We developed test script in order to test Wysa app, which is a chatbot that offer mental health counseling
- Our final version of test script supports
  1. Multiple test cases within one run
  2. The test script reads multiple input and expected output from Excel file
  3. Has similarity string comparison using Sorensen-Dice coefficient

## Technologies:

- The test script used Javascript because the language offer good asynchronous method
- Runs on Android simulator
- We used Appium as testing server

## 3. Test scripts structure

1. In order to perform multiple test cases with just one run, our test scripts read all the input and expected output from Excel file, as shown in Figure 2.1

| Input   | Expected output   |
|---|---|
| Give me some mindfulness exercises for Asian elders                         | Would you like to try some exercises right now?         |
| I want to set some exercise routines as an Latino adult                     | Let's set up exercise routine together                  |
| This exercise was helpful for me  | Thanks for your feedback                                |
| This exercise was not helpful for me  | I am sorry I will strive to do better                   |
| I want to reframe thoughts to feel better as a Latino kid                   | Would you like to try some reframe thoughts right now?  |
| As a Latino kid, I want to set schedule to reframe thoughts every Wednesday | Let's set up reframe thoughts schedule                  |
| The reframe thought exercise was helpful                                    | Thanks for your feedback                                |
| The reframe thought exercise was not helpful                                | I am sorry I will strive to do better                   |
| Give me some strength exercises for Asian elders                            | Would you like to try some exercises right now?         |
| I want to set some strength exercise routines as an Asian elder             | Would you like to try some exercise routines right now? |
| This strength exercise was helpful for me                                   | Thanks for your feedback                                |
| This strength exercise was not helpful for me                               | I am sorry I will strive to do better                   |
| Give me some durability exercises for unemployed adult                      | Would you like to try some exercises right now?         |
| I want to set routine to domindfulness every Monday as anLatino adult       | Would you like to set up some routines right now?       |
| This mindfulness exercise was helpful for me                                | Thanks for your feedback                                |
| This mindfulness exercise was not helpful for me                            | I am sorry I will strive to do better                   |
| Tôi muốn tập thể dục  | I can only respond in English                           |
| I want to do physical exercises instead                                     | Let's change the topic for now                          |

Figure 2.1 Input and Expected output of an Excel file

2. The test script will then enter the input, collects the output, as shown in Figure 2.2
3. Then it compares the output with expected output using Sorenson-Dice coefficient
4. If the similarity score passes a certain threshold, then the test cases considers passed

The screenshot shows a Node.js development environment with the following details:

- EXPLORER:** Shows files in the project directory, including `AutomateTestCases.js`, `ExerciseFunction.js`, `FeelingDetectionFunction.js`, `InvalidDetectionFunction.js`, `MentalIllnessFunction.js`, `read.js`, `temp.xlsx`, and `testResults.txt`.
- Code Editor:** The `AutomateTestCases.js` file contains the following code:

```

    1 // Requires the webdriverio client library
    2 // (npm install webdriverio)
    3 // Then paste this into a .js file and run with Node:
    4 // node <file>.js
    5
    6 const similarity = require('string-similarity-js')
    7 const wdio = require('webdriverio');
    8 const fs = require('fs')
    9 const reader = require('xlsx')
    10
    11 async function automateTest() {
    12     const file = reader.readFile('./temp.xlsx')
    13
    14     let data = []
    15     let indexStarts = 8
    16
    17     similarityPassedThreshold = 0.25
    18
  
```

- TERMINAL:** Shows log output from the WebDriver session, including commands like `elementClear`, `elementSendKeys`, and `findElement`, along with INFO messages about file reads and POST requests.
- Android Emulator:** Displays a chatbot interface with the following conversation:
  - User: "No, you misunderstood"
  - Bot: "Oh! I'm sorry"
  - User: "Shall we go back to what we were talking about?"
  - Bot: "Something else?"
  - User: "What would you like to try now, Hmmm?"
  - Bot: "This strength exercise was not helpful for me"
  - User: "I sense you are finding this conversation unhelpful. Is that right, Hmmm?"
  - Bot: "Yes" (with a yellow thumbs-up icon)
  - User: "Not really, go on" (with a yellow thumbs-down icon)
  - Bot: "Give me some durability exercises for unemployed adult"

*Figure 2.2 The test script is being performed*

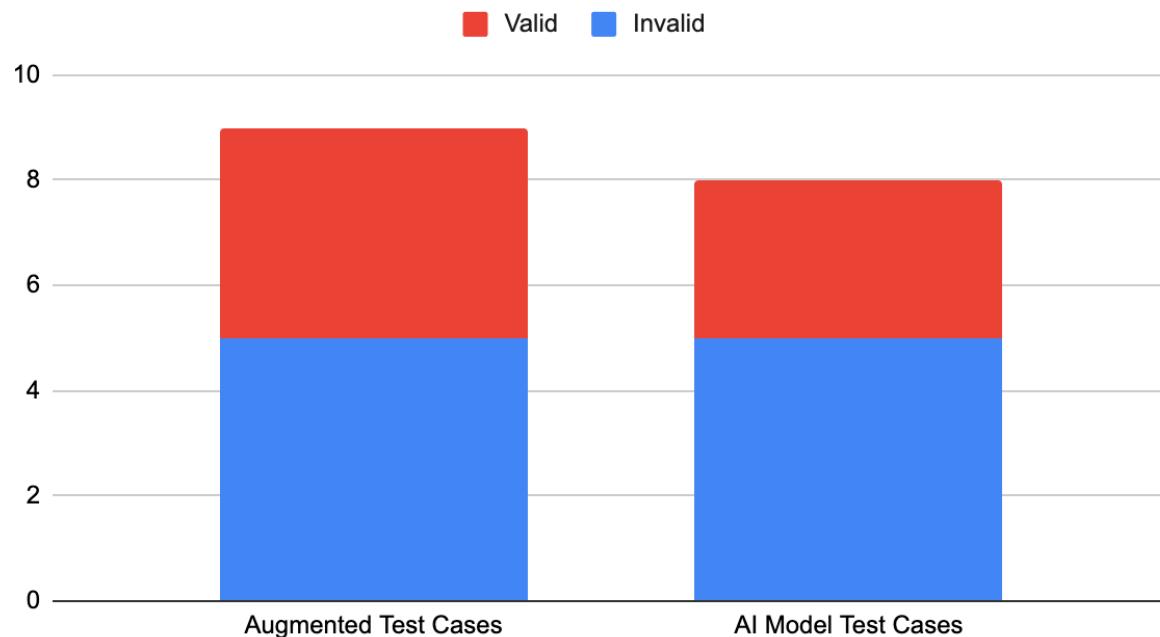
## Test automation results

### Test result

#### Mental Illness function

- For this function, we automated a total of 17 test cases
  - We tested 9 data augmented test cases
    - 5 out of 9 test cases failed,
    - 44.44% success rate, 55.56% failure rate
  - We tested 8 AI powered test cases
    - 5 out of 8 test cases failed
    - 37.5% success rate, 62.5% failure rate

#### Mental Illness Test Case Analysis



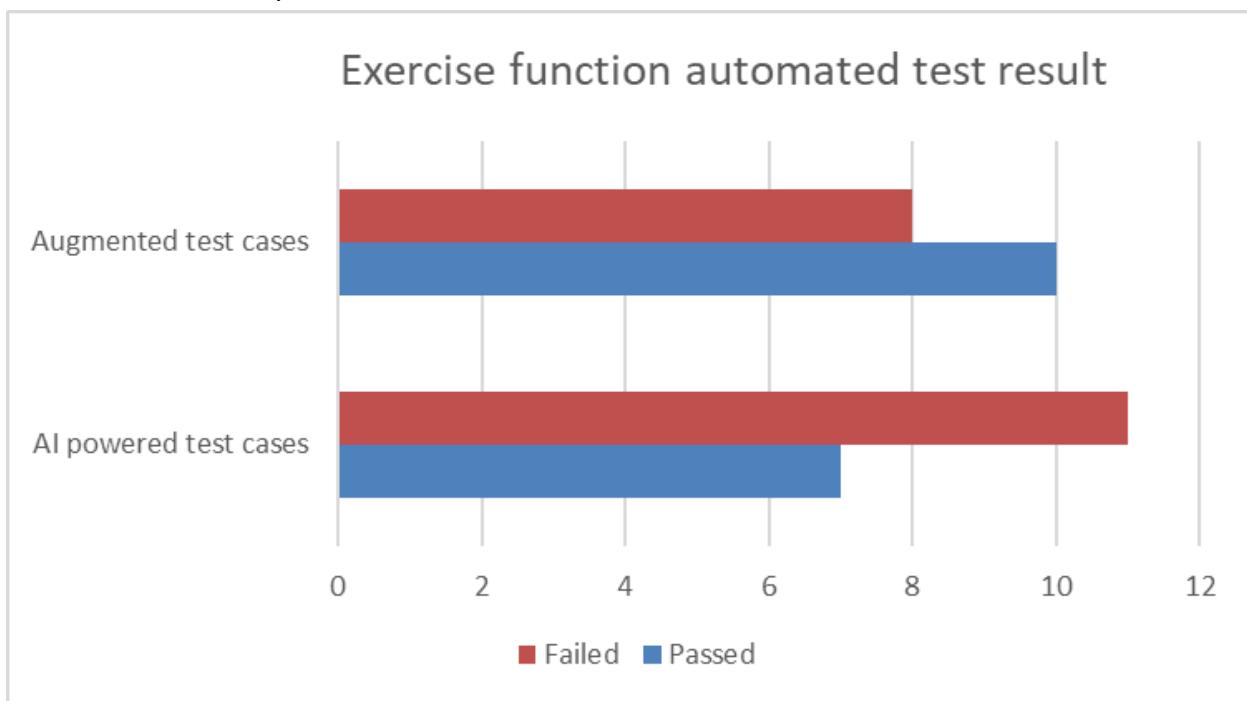
#### Exercise function

- For this functionality, we automated 36 test cases
  - 18 of them are AI powered test cases
    - At least one test case cover for each minor case in the input and context spanning trees
  - 18 augmented test cases
    - Random word swap

- Random word delete
- Random word insert
- Synonym augmentation
- Ocr augmentation
- Keyboard augmentation
- Random char insert
- Random char swap
- Random char delete

Result is as below

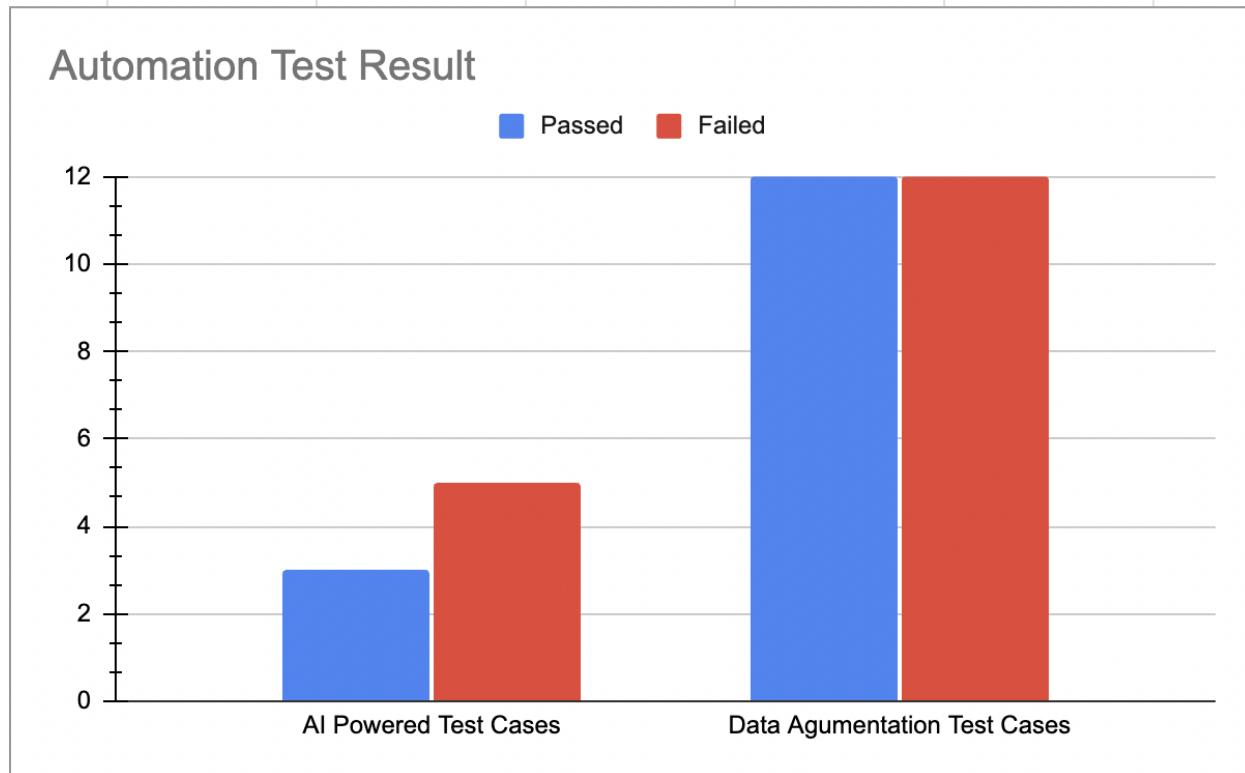
- AI powered test cases:
  - 7/18 passed, 38.89%
- Augmented test cases:
  - 10/18 passed, 55.56%



## Feeling Detection Function

- We automated 32 test cases for this function
  - 8 of which are AI powered test cases
    - At least one test case cover for each minor case in the input and context spanning trees
  - 24 of which are data augmentation test cases
    - Random word insert
    - Random word delete
    - Random word swap
- Automation Test Result
  - AI powered test cases

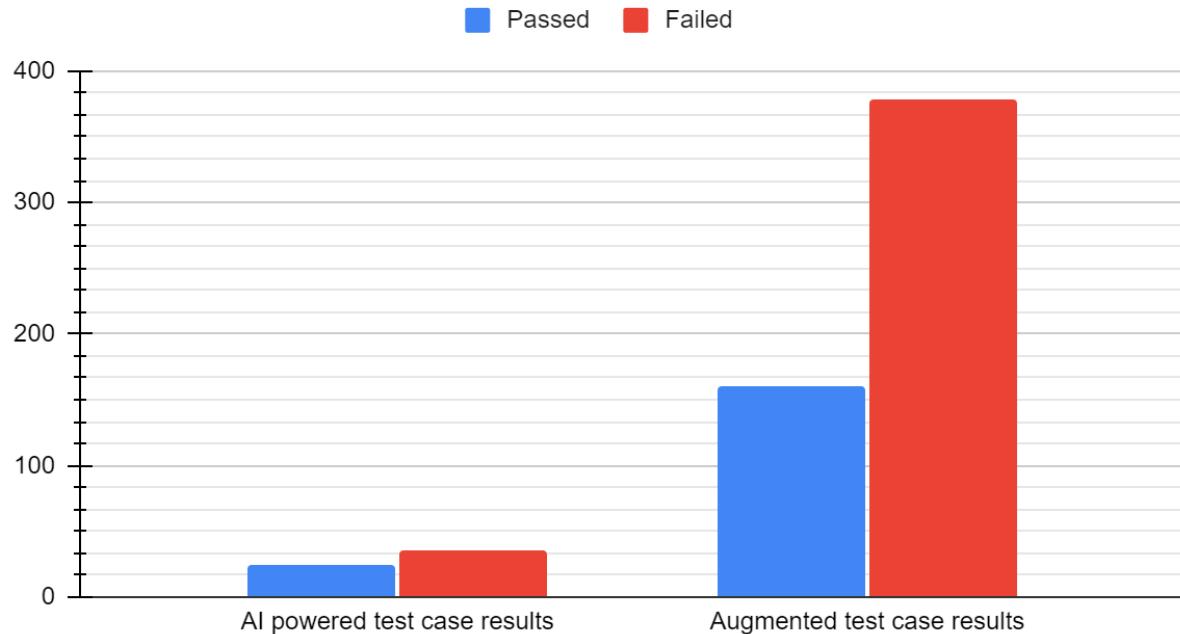
- 3/8 passed = 37.5%
- Data augmentation test cases
  - 12/24 passed = 50%



## Invalid Topics

- This function automated 600 test cases
  - 60 of these are AI powered test cases
    - At least one test case for every input and context tree combo
  - 540 are augmented test cases
    - Random word swap
    - Random word delete
    - Random word insert
    - Synonym augmentation
    - Ocr augmentation
    - Keyboard augmentation
    - Random char insert
    - Random char swap
    - Random char delete
- Automated results
  - AI powered test cases
    - 24/60 passed = 40%
  - Augmented test cases
    - 161/540 passed = 29.8%

## Invalid Topics Automated Test Result



## Test cost

### Mental Illness function

- Each test takes about 20 seconds to complete
- The total time spent to test this function is  $20 \times 17 = 360$  seconds, or 6 minutes

### Exercise function

- Each test takes about 25 to 30 seconds to complete
- Therefore with 36 test cases, we spent a total of  $30 \times 36 = 1080$  seconds, or 18 minutes

### Feeling Detection Function

- Each test case take about 20 to 30 seconds to complete
- We spent about 14 minutes for this function

### Invalid Topics

- Each test takes roughly 10 to 25 seconds to complete
- In total an average of 175 minutes, or almost 3 hours, is spent to test this function

## Test complexity

| Test function              | Test complexity |
|----------------------------|-----------------|
| Mental Illness function    | 17              |
| Exercise function          | 24              |
| Feeling detection function | 32              |
| Invalid topics             | 600             |