**Diagram, text

Description automatically generated**

**Introduction:**

During the recent years, many different technologies were introduced in the field of radiology, those technologies have been developed and evolved to be used in broad applications. Voice and speech recognition technology is considered as the fastest growing artificial intelligence technology in this field, it is used to many different applications and systems such as voice interaction, live streaming, traffic, and home automation by companies. This is also known as virtual assistant A.I technology.

Virtual assistant usually uses cloud computing such as Amazon Web Services, Siemens Mindsphere, and Microsoft Azure to save the programs, where it requires internet connected devices to work, where the main goal for virtual assistants is to interact with users to answer questions and undertake different tasks, such systems are Apple’s Siri, Amazon Alexa, and Google Assistant [1].

There are many factors that are needed in natural language processing to design a successful virtual assistant, such as [2]:

* Speech-to-text
* Text-to-speech
* Syntax and Semantics processing
* Question Answering

Although using the four factors above a successful virtual assistant can be designed, there are many different situations where the assistant will struggle to give the ideal output that is expected by the users. In order to know what those situations are, different test cases are applied on the system, where each test case can include different inputs and different environments.

This paper will be mainly focusing on the testing of Amazon’s Alexa virtual assistant by looking at the particular issues involved in evaluating the system. The evaluation was undertaken by applying different test cases in order to understand the user interaction with Alexa.

Specific Alexa skills and capabilities are chosen for the testing of the system, the chosen skills to test are:

* *Amazon’s Alexa capability of understanding speech in different accents and environments.*
* *Amazon’s Alexa capability of playing music from different music platforms.*

In some cases, the same input will be given to Alexa but with different accents, language, or in a different environment, noticing the output will also be different in some test cases.

**Methodology:**

All test cases conducted in this paper follow the Black Box method of testing.

In order to examine the mentioned Alexa skills, many test cases can be applied with different inputs to the system. First, focusing on the first Alexa capability:

*Amazon’s Alexa capability of understanding speech in different accents and environments.*

The rationale of choosing the test cases for the capability above is because the users that use Alexa are from all over the world, hence, different accents are being used when the input commands are given to Alexa. The goal of this evaluation is to test if Alexa will understand the same command given different accents and environments.

Alexa’s settings are controlled through an app that can be installed on iOS and Android, through the app there is an option to choose different languages depending on the preference of the users. A test case as a foreigner trying to speak a specific language would be reading a sentence in a language when it is not the user’s first language and evaluate if the language processing would understand it and give an ideal output. Depending on the nationality of the users and the language they speak, some languages dialogues are faster or slower than others, which might affect Alexa from understanding the command, therefore, Alexa is tested by giving input commands when speaking fast and slow than usual. Thus, the expectation is that one of Alexa's requirement would have been to be able to understand many accents.

Another challenge for virtual assistants in general is usually the environment where the user is located while giving the input, for example, a very noisy environment while giving an input command may result in Alexa misinterpreting the command and not responding to it, the rationale choosing a test case for this is because users might want to command Alexa to do something during a gathering with loud people and music, therefore, it is an important capability to test.

Alexa uses voice recognition, meaning that it can be trained to know who is talking to it and giving it commands in real time, meaning when the Alexa owner asks, “Who is talking?”, Alexa should respond by the name of the person talking and should be able to distinguish between different voices, different test cases are built over this topic, the rationale for the test cases on this topic is because Alexa is usually used in a shared home which includes various users, each of the users can have their own to do list for example, therefore Alexa needs to know who is talking in order to give the correct output to the correct user [3].

As Alexa is usually located in a house or an office, the users might need to give their input command from different rooms, it is important to know if Alexa will be able to understand the given speech from another room and if it will be able to give a valid output. Therefore, different test cases for that situation are also evaluated where different outputs and results are given depending on many factors such as the distance, the volume of the voice, and the location of Alexa in the room.

As this paper is focusing on two different Alexa capabilities, it is time to discuss the second Alexa capability and the rationale of the test cases.

*Amazon’s Alexa capability of playing music from different music platforms.*

The Alexa has the possibility to play music when it is commanded to do so by the user from different music platforms such as Spotify, Amazon Prime Music, and from many different radio stations.

In order to test the music playing capability in Alexa, different inputs are given depending on the preferences of the user. There are many conditions under which Alexa’s music playing capability can be tested. The goal is to create real life conditions for testing the capability.

* If a user has just bought Alexa and he commanded it to play music, what is the default music platform that Alexa will choose?
* Asking Alexa for the lyrics of a specific song.
* Asking Alexa to play music from a platform where the user doesn’t have an account?
* Asking Alexa to play an Arabic-named song from Spotify.

All of those are situations where the music capability test cases are tested, other situations include if Alexa is playing music on maximum volume, the user wants to give it a new command, will Alexa listen to the command from the music? This is an environmental issue where the surroundings affect the performance of the system.

As all the music platforms include songs from all around the world where the titles of the songs are in different languages, if the user is Arabic and he wants to ask Alexa to play a specific Arabic song for him (Alexa does not have Arabic language in its setting), will Alexa be able to understand the Arabic name of the song and play it from a specific music platform? This situation is could also be added under the speech recognition factors.

In order to test the understanding and the theory behind the machine learning in Alexa in this topic, as testers, it is important to know if Alexa understands the user’s feelings, what if the user tells Alexa, “Alexa, I’m sad, play me some music”. What would be Alexa’s response and what music genre will Alexa play?

Many different test cases were undertaken in order to evaluate this specific capability. The next section in this report will be discussing the results of the test cases in general in both the capabilities discussed in the methodology. All the test cases that were evaluated on both capabilities are chosen to important rationales of real-world inputs that the user of Alexa would try.

**Results:**

Many different test cases to evaluate the system were tackled in order to test the speech recognition and the music capabilities of the Amazon Alexa assistant, although there were some cases were some expected results which were not given as an output, adding it to the “Fail” list, however, in some other cases the expected result was given by Alexa and therefore, it is added to the “Pass” list for this test case.

Regarding the understanding of speech in different languages and accents as well as loud environments, 14 test cases were handled in this area, such as speaking in a different language, evaluating if Alexa knows who’s talking, speaking slowly and checking if Alexa can understand, and many more.

Out of the 14 test cases done in this area, Alexa has passed only 6 cases and failed 9 cases. In many of those test cases Alexa did not give any respond, or it just said, “I do not understand the question”.

Other test cases were undertaken to evaluate the *music capability* in Alexa, this included test cases on all the aspects related to music in the system, such as changing volume, shuffling music, playing from different platforms, and many more. In this area, 21 test cases were made to evaluate this skill, Alexa has only passed in 40% of the test cases, meaning it only passed 9 test cases giving the expected output and failed 13 other test cases where the output was different than the expected one.

Therefore, a total of 35 test cases were used to evaluate the system in two different capabilities / skills, the system has passed 15 cases and failed 20 test cases. This is based on different bugs in Alexa and shows how in some situations it is affected by the surrounding environment.

|  |  |  |
| --- | --- | --- |
| Total Test Cases | 35 | 100% |
| Passed Test Cases | **15** | **42%** |
| Failed Test Cases | **20** | **57%** |

**Conclusion:**

Many virtual assistants are in-use by different users all around every day, where users give unexpected inputs the virtual assistant. As a virtual assistant, it learns more about the users input when more commands are given to it, it is always trained and improved by different data. Virtual assistants make it easier for users to apply many applications such as alarms, reminders, smart homes control, and playing music. As the main virtual assistant in this paper is Amazon’s virtual assistant “Alexa”, it has been tested with 35 different test cases to test its abilities in voice recognition and understanding, as well as playing music from different streaming platforms. Alexa has failed around 55% of the test cases, which is a large value for a system that is used worldwide.

There are many different ways in which the users can help decrease the amount of failed test cases, such as giving the inputs in different forms and tell Alexa when she did not give the correct output, as it is an A.I system, Alexa will be trained by new data and improve itself. Another way to help decrease the failed situations is to log in to the Amazon website, all you inputs are saved and displayed in your account, where it asks you if the answer for this input was correct, this can help the Amazon engineers to improve Alexa as they understand where some of the issues are.

**References:**

[1] European Commission. 2018. “The rise of Virtual Personal assistants”. [Online]. Available: <https://ec.europa.eu/growth/tools-databases/dem/monitor/sites/default/files/Virtual%20personal%20assistants_v1.pdf> (Accessed 25 October 2020)

[2] I. Lopatovska, K. Raines. 2018. “Talk to me: Exploring user interactions with amazon alexa”. [Online]. Available: <https://www.researchgate.net/publication/323635232_Talk_to_me_Exploring_user_interactions_with_the_Amazon_Alexa> (Accessed 27 October 2020)

[3] M.B.Hoy. (2018). “Alexa, siri, cortana, and more: An introduction to voice assistants”. [Online]. Available: <https://www.researchgate.net/publication/322456429_Alexa_Siri_Cortana_and_More_An_Introduction_to_Voice_Assistants> (accessed 29 October 2020)

**APPENDIX:**

Included in the appendix, two tables, where the first table includes the test cases for the first Alexa capability which is understanding speech by different accents and in different environments.

The second table includes all the test cases for the second capability which is the Music Playing capability.

Please a ZIP File included in the submission (and privately emailed) which includes all the audio recordings of all the test cases below.

1. ***Tests cases for the first capability:***

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Test Cases Number | ***Test*** | ***Utterance*** | ***Expected Result*** | ***Pass / Fail*** |
| 1 | Yassin asks Alexa who is talking. | “Alexa, do you know who is talking?” | Alexa knows that Yassin is talking | Pass |
| 2 | Someone else asks Alexa who is talking. | “Alexa, do you know who is talking?” (Another person) | Alexa should not know who is talking | Pass |
| 3 | Asking Alexa is Spanish who is the queen of Britain (Foreigner Accent) | “Alexa, que es la reina de Gran Bretaña” | Alexa should understand my Accent and answer. | Fail |
|  |  |  |
| 4 | Ask Alexa if she can hear me while there is music in the background. | “Alexa, can you hear me?” | Alexa respond by “Yes, I am here” | Pass |
| 5 | Ask Alexa from a different room about the time” | “Alexa, what’s the time” | Alexa should tell me the current time | Pass |
| 6 | Ask Alexa about the weather while speaking very slowly. | “Alexa, what is the weather?” | Alexa should tell me what the weather is. | Fail |
| 7 | Ask Alexa about the weather while speaking very fast. | “Alexa, what is the weather”? | Alexa should tell me what the weather is. | Fail |
| 8 | Asking Alexa about a football game while whispering | “Alexa, what time is the Barcelona game tonight? | Alexa should tell me the time of the game | Pass |
| 9 | Asking Alexa to tell me a joke, then interrupting the joke and ask Alexa about the weather | “Alexa can you please tell me a joke”? – “Alexa what is the weather”? (Interrupting the system) | Alexa should start telling me the joke, should stop the joke when I call it and follow the next command. | Fail |
| 10 | Ask Alexa to turn the volume down while it is playing music. | “Alexa, please turn the volume down” | Alexa should turn the volume down. | Fail |
| 11 | A 2 years old child asking Alexa about the location of Antarctica | “Alexa, where is Antarctica” | Alexa should understand the child’s accent and respond to the question | Fail |
| 12 | Stopping the Alexa morning alarm from another room. | “Alexa, stop the alarm” | Alexa should stop the alarm. | Pass |
| 13 | Ask Alexa if I can be heard when there is a loud football game on the TV (Background Noise) | “Alexa can you hear me?” | Alexa should respond and say yes. | Fail |
| 14 | Record my voice while asking Alexa who is talking and play the recording to Alexa. | "Alexa, do you know who is speaking?" (Digital Voice) | Alexa should say Yassin is talking | Fail |

1. ***Tests cases for the second capability:***

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Test Cases Number | ***Test*** | ***Utterance*** | ***Expected Result*** | ***Pass / Fail*** |
| 15 | Asking Alexa to play music when no music platform accounts are linked | “Alexa, play music” | Alexa should play music from a default platform (Amazon Prime Music) | Pass |
| 16 | Asking Alexa to play an Arabic song (with Arabic name) from Spotify | “Alexa, play “Bent El Giran” by Hassan Shakosh from Spotify” | Alexa should understand and play the song from Spotify. | Fail |
| 17 | Ask Alexa to play a song, then add another song to the queue. | “Alexa please play When to Say When by Drake and then add Yesterday by the Beatles to the queue” | Alexa should play both songs back to back. | Fail |
| 18 | Tell Alexa my mood to listen to music. | “Alexa, I want to dance, play music” | Alexa should play dancing music | Pass |
| 19 | Ask Alexa to play music similar to Frank Sinatra’s. | “Alexa, play something similar to music by Frank Sinatra” | Alexa should play similar Blues music. | Fail |
| 20 | Ask Alexa to add a song my playlist. | “Alexa, add this song to my Jazz playlist on Spotify” | Alexa should add the song to my playlist and confirm it. | Fail |
| 21 | Ask Alexa to create a new playlist and give it a name. | “Alexa, create a playlist on my account and call it Love Music” | Alexa should create the playlist and confirm it. | Fail |
| 22 | Ask Alexa to play music from a radio station | “Alexa please play music from a radio station” | Alexa should choose any radio station and play music. | Pass |
| 23 | Tell Alexa I don’t like the current song and ask Alexa to change it. | “Alexa, I don’t like this song, please change it” | Alexa should stop the current song and play the next one. | Fail |
| 24 | Ask Alexa to repeat a song that is being played. | “Alexa, can you please repeat this song” | Alexa should reply the song from the beginning. | Fail |
| 25 | Ask Alexa to play popular music from Spotify. | “Alexa, can you please play popular music? | Alexa should play music from the top charts. | Fail |
| 26 | Ask Alexa to shuffle the current playlist. | “Alexa, please shuffle this playlist” | Alexa should shuffle the playlist and put music in random order. | Pass |
| 27 | Ask Alexa to play a specific music genre | “Alexa, please play Country music” | Alexa should play country music from any platform. | Pass |
| 28 | Ask Alexa to play a song I haven’t heard in a while | “Alexa, please play a song I haven’t heard in a while” | Alexa should play a song that haven’t been played on my Spotify account for a long time. | Pass |
| 29 | Ask Alexa to play a song by giving the lyrics as input | “Alexa, play a song that goes “We are not strangers to love”. | Alexa should play a song that has those exact lyrics. | Fail |
| 30 | Ask Alexa to give the lyrics for a song. | “Alexa, can you please tell me the lyrics for the song Savage Love by Jason Derulo”? | Alexa should tell me the lyrics of the song. | Pass |
| 31 | Ask Alexa to play music from a music platform that is not pre-registered. | “Alexa, can you please play Music from Deezer? | Alexa should ask me to register to this platform in order to play music from it. | Pass |
| 32 | Ask Alexa to play a specific genre from another streaming platform. | “Alexa, can you please play Music from Deezer? | Alexa should not play music as platform has no linked account. | Pass |
| 33 | Asking Alexa who is the king of pop then asking Alexa to play music by the King of Pop | "Alexa, who is the king of pop?" -- "Alexa, play music by the king of pop. | Alexa should say the King of Pop is Michael Jackson and play his music. | Fail |
|  |  |  |  |  |
| 34 | Asking Alexa what is my default music platform after changing it to Spotify. | "Alexa, what is my default music platform?" | Alexa should say that Spotify is the default music platform | Fail |
| 35 | Asking Alexa for good music recommendation | "Alexa can you tell me good music recommendation" | Alexa should tell me good music recommendation | Fail |