[Date]

ishtiyaq ali

[company name]

[Document title]

[Document subtitle]

Abstract

Small paragraph…..

Contents page

# **Introduction**

## **Purpose**

In the area of speech recognition, it is said that Microsoft can now interpret *human speech with a 5.1% error rate*. Google have enhanced its *accuracy by more than 20% in the past five years*. And to date, Amazon's Alexa has been getting better at responding to users’ question. *Researchers asked the voice assistant 800 different queries* in various categories. *On average, Alexa answered queries accurately 73% of the time, up 12 percentage points from 61%*. These are little known facts of how top companies are taking speech recognition to a whole new level and how they've made it worth looking into and investing in for personal and business use. Given that the unexpected rise has happened so recently, this project will aim to contribute to the area by researching the tools and techniques that are being used to make it work so efficiently. The research will then be used to implement an Android application called Speech Radar, that will allow users to locate their phone through speech recognition. Suppose for instance an individual must rush to work and is unable to find their phone, but knows it’s located somewhere inside their room. Speech Radar can speed up the search time just by an utterance of a specific word from the individual. Once the phone detects the utterance, it will begin ringing at maximum sound level, which will help the individual locate their phone. This app can be of use to a great number of people as it is common for individuals nowadays to lose sight of their phone and to search everywhere for it.

## **Background**

Since 1784, Speech recognition is something that was just a topic of talk. It wasn't until 1952 when a six-foot machine was created by Bell Labs, *capable of recognizing spoken digits with 90% accuracy*, but when uttered by its owner. The development would continue in 1962, where IBM created a machine the size of a shoebox that could *understand 16 English words*. In 1971 a student of Carnegie Mellon University created the Harpy that could *comprehend 1011 words and some phrases*. In 1986, IBM would create another ground-breaking machine that used the *Hidden Markov Model* to recognise 20000 different words from various speakers and type them on paper. The list of inventions would go on with Google launching a voice search application in 2008, *bringing speech recognition to mobile devices*. In 2011, Apple would announce Siri, *ushering in the age of the voice-enabled digital assistant.*

At present, we are seeing digital assistants decentralise from smartphones and are seeing companies primarily focus on voice-activated home speakers that can query and control smart home devices. From a subjective point of view, these innovations appear to be an approach to accumulate billions of audio data from people that have different accents, so companies in the future can improve the detection rate for fluent and non-fluent English speakers. On the off chance that this improvement happens, we will see speech recognition being used for more advanced tasks, possibly in robotics.

## **Aims and objectives**

* To research on a feasible approach of creating a TensorFlow model for speech recognition by going through implementations of certain users public GitHub repositories
* To research on an appropriate dataset to use for training the model. Preferably, the dataset should contain audio files for each word in different accents, rather than phrases.
* To design our implementation in TensorFlow and Android Studio using pseudocode, flowcharts, wireframes and UML diagrams
* To implement our designs and test the speech recognitions detection rate with many random users. Gather the results for analysis, and use it to improve the application

## **Section overview**

Chapter 2

Chapter 3

Chapter 4

Chapter 5

Chapter 6

Chapter 7

Chapter 8

Chapter 9

# **Literature review**

# **Specification**

# **Design and implementation**

# **Testing and evaluation**

# **Conclusions and future work**

# **Appendices**

# **Reference list/bibliography**

<https://www.temi.com/blog/little-known-facts-about-speech-recognition-technology/>

<https://www.businessinsider.com/amazon-improving-alexa-voice-accuracy-2018-12?r=US&IR=T>

<https://medium.com/swlh/the-past-present-and-future-of-speech-recognition-technology-cf13c179aaf>

# **Evaluation**