AUTOMATED DOCTOR APPOINTMENT MANAGEMENT SYSTEM FOR MEDICAL DOMAIN IN SINHALA USER(ADAMS).

2020-175

Project Proposal Report

K.N.B. Kudawithana

B.Sc. (Hons) Degree in Information Technology

Department of Information Technology

Sri Lanka Institute of Information Technology
Sri Lanka

March 2020

APPOINTMENT MANAGEMENT SYSTEM

2020-175

Project Proposal Report

B.Sc. (Hons) Degree in Information Technology

Department of Information Technology

Sri Lanka Institute of Information Technology
Sri Lanka

March 2020

DECLARATION

We declare that this is our own work and this proposal does not incorporate without acknowledgement of any material previously submitted for a degree or diploma in any other university or Institute of higher learning and to the best of our knowledge and belief it does not contain any material previously published or written by another person except where the acknowledgement is made in the text.

Name	IT Number	Signature	Date
Kudawithana K.N.B	IT17255820		25/02/2020

Supervisor

---------- ----------

Prof. Koliya Pulasinghe Date

Co-Supervisor

---------- ----------

Ms. Vijani Piyawardana Date

ABSTRACT

Communication is essential for the evolution and development of all kinds of living beings. Without discussion, languages must be recognized as the most amazing artefacts that humanity has developed to allow communication. The computer has also become such a unique mass due to its ability to communicate with people through language. It is worth noting that the languages that computers and people understand are quite different, but people can communicate with a computer. This was possible because the computer is, in principle, an artefact that can translate one language into another. Therefore, computers should be able to do language translations than any other computer task. Today, computer science is evolving to allow machine-to-machine communication without any small human intervention, but people still face what is called a language barrier to communication. In particular, a vast collection of world knowledge written in English was inaccessible to communities that could not communicate in English. However, due to the inherent diversity of natural languages, a generic approach to machine translation is far from reality.

But all of these systems are for computers and there are very few communication systems in Sinhala using mobile phones. The thing is there are no any automated system in Sinhala for make an appointment and automatic disease identification using voice input. This mobile application called as ADAMS. It is expert system for identify diseases from given symptoms and user can make an appointment easy using Sinhala as language without type any word.

From this part it will be translate given Sinhala sentences to English sentences, identify given sentences with (NLP,) Search and divide medical terms and facts from the sentences After all generate query for them are the tasks. For collect data about medical centers there should have implemented software so to complete this task ADAMS will have another sub software with flutter.

Keywords -: Firebase, Flutter, NoSQL, NLP, English to Sinhala Translate

LIST OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

1. INTRODUCTION

Natural Language Processing (NLP) is a combination of artificial intelligence and linguistics designed to make computers understand statements or words written in human languages. Natural language processing has appeared to facilitate the user's work and satisfy the desire to communicate with a computer in a natural language. Since not all users can be well versed in a machine-oriented language, NLP serves those users who do not have enough time to learn new languages or improve on them. In this study, people who do not speak English and people with disabilities will find it easier to schedule an appointment with a doctor.

Given Sinhala sentences will be translating in to English, English sentences will be processed and identify medical facts comparing with database and generate them to automated query for TTS process. There will be another flutter application for register medical centres.

1.1 BACKGROUND & LITERATURE SURVEY

NLP is a subfield of computer science and artificial intelligence concerned with interactions between computers and human (natural) languages. It is used to apply **machine learning** algorithms to **text** and **speech**.

For example, can use NLP to create systems like **named entity recognition and Sentence Recognition** are the parts of this research. Nowadays, most of us have smartphones that have speech recognition. These smartphones use NLP to understand what is said. Also, many people use laptops which operating system has a built-in speech recognition**.**

In here for separate important medical domain can use gradient algorithm.

Computing the solution of the optimization objective is achieved by minimizing the cost function. Its gradient, denoted by ∇C, reveals in which direction the function decreases the fastest. Utilizing ∇C it is possible to repeatedly step towards a local minimum.

The learning rate α is a hyper parameter and is manually tuned. There are methods that approximates a learning rate [7] speeding up convergence, but here Gradient descent is kept in a pure form. When α is chosen far from its optimal value $\alpha *$ the algorithm will diverge if $\alpha > \alpha*$and converge very slowly if $\alpha < \alpha*$. [1]

Recall the optimization objective C(θ) in. Let

$$h_{\boldsymbol{\theta}}(\mathbf{x}) = \boldsymbol{\theta}^T \mathbf{x}$$

FIGURE 1 : GRADIENT ALGORITHM – FORMULA 1

If the number of training examples is m, the number of parameters j = 1, ..., p and e1, ..., ep is the standard basis, then

Where

$$\nabla C(\boldsymbol{\theta}) = \frac{\partial C(\boldsymbol{\theta})}{\partial \theta_1}\mathbf{e}_1 + ... + \frac{\partial C(\boldsymbol{\theta})}{\partial \theta_p}\mathbf{e}_p$$

$$\frac{\partial C(\boldsymbol{\theta})}{\partial \theta_j} = \frac{1}{m}\sum_{i=1}^{m}(\boldsymbol{\theta}^T \mathbf{x}^{(i)} - y^{(i)})\mathbf{x}_j^{(i)}$$

FIGURE 2 : GRADIENT ALGORITHM – FORMULA 1

The pseudocode for gradient descent is presented in algorithm 1. The initial θ0 is chosen arbitrarily.

Algorithm 1 Gradient descent

1: Initialize $\boldsymbol{\theta}_0$
2: Set iterations L
3: Set learning rate α
4: **for** k = 0 to L **do**
5: Evaluate $g_k = \nabla C(\boldsymbol{\theta}_k)$
6: $\boldsymbol{\theta}_{k+1} \leftarrow \boldsymbol{\theta}_k - \alpha g_k$
7: **end for**

FIGURE 3 : GRADIENT ALGORITHM

Convergence of gradient descent is dependent on choosing an accurate learning rate, as well as enough iterations. But given a convex optimization problem as in and appropriate hyper parameters, i.e. $0 < \alpha < 1$, the algorithm will converge. [1]

In, which is the optimization objective for logistic regression, another cost function is defined. Given the structure of gradient descent, it is apparent why J has been defined in such a way.

Consider,

$$\frac{\partial J(\boldsymbol{\theta})}{\partial \theta_j} = \frac{1}{m} \sum_{i=1}^{m} \mathbf{x}^{(i)} (h_\theta(\mathbf{x}^{(i)}) - y^{(i)})$$

Figure 4 : Gradient algorithm – formula 3

The gradient of this function can be computed as in. Furthermore, since J is a convex, gradient based algorithm can be applied to minimize. Formally, this is a logistic regression problem and the model in section 4 will be weighted using. [1]

Generally, the literature agrees that NoSQL databases possess greater flexibility and scalability than traditional SQL databases but at the expense of functions that are taken for granted in relational technologies. Consequently, much of the research surrounding NoSQL concerns the classification of NoSQL technologies, the relative advantages and disadvantages of each category, how shortcomings may be remedied (especially with regards to consistency, querying and interoperability), and the adoption of NoSQL technologies.

1.2 RESEARCH GAP

Using Google API's and Gradient Algorithm already done some of features which are mentioned below. Those features implemented in a desktop application in Sinhala language. Through ADAMS done all the developments under mobile application using RASA framework in medical domain.

None of these applications didn't able to develop the Sinhala terms dictionary to gather Sinhala medical terms knowledge. Also didn't able to display whole conversation on an interface.

Features	Google API (desktop version)	Laliga Chat Bot (desktop version)	RASA Framework on Flutter mobile application and Gradient Algorithm
Sinhala to English Medical domain convertor	✔	✘	✘
Identify Sentence (English & Sinhala)	✔	✔	✔
Separate medical facts (English)	✔	✔	✔
Generate query for selected medical facts (English)	✘	✘	✔
Make firebase database to connect flutter application	✔	✘	✘
Flutter application for collect data about medical centers	✘	✘	✘

TABLE 1 : RESEARCH GAP

2. RESEARCH PROBLEM

Nowadays in Sri Lanka, all of the e-channelling systems are based on the English language. Mostly e-channelling systems are used by people in western province comparing to other provinces. The reason is the erudition that they have is very impecunious in the English language.

Sometimes, people having diseases such as Short-Term and Long-Term Incapacitation and People with dyslexia, are incapable of using communicate an E-channel system.

Most of the Sri Lankan e-channelling systems are web-predicated and utilizing the web-predicated system in mobile is Minimize Celerity, sometimes browser support is very impuissant and not facile to utilize on mobile. It is not user-friendly for mobile.

Most of the patients don't know what the specialization of the doctors is, and who are the best doctors of their diseases. They only know the diseases that they have at that moment.

Some of the patients know doctor's name but they will not know the details about the doctor, such as the hospital where, the doctor is available, time schedules of the doctors.

Some patients know all the details and they want to get an appointment with doctors, but they couldn't make it, because they will face interaction problems with a system, the system will perform the English language they can't understand it.

Patient have to pace on above difficulties in this kind of a challenging situation.

3. OBJECTIVES

Create Firebase Database

Firebase is a trending Cloud based (BaaS) backend as a service which provides Real Time syncing with all the clients subscribed to the server at any given instance. Firebase supports multiple platforms/frameworks like AngularJS, Backbone.js, iOS7, OSX, Android and programming languages like Ruby, Node.js, Python, Java, PHP, Perl & JavaScript. Firebase relies on web sockets to update all the clients about any data changes instantly.so ADAMS can use Firebase as Database.

Translate Given Sinhala and English Sentence

The Translator is used to translate an English base word into a Sinhala base word with the help of the bilingual dictionary. The translator translates subject, object and verb in the English sentence separately. This is a method we have used to reduce the complexity of the translation process. This translator is a simple one and it does not automatically handle the semantics of sentences. Note that, this stage can be supported by human intervention to generate the most appropriate translation for some words in a sentence. As such, handling semantic, pragmatic and Multiword expressions must be addressed with the support of humans, for which we introduce an intermediate editor.

Divide medical facts from given sentences

In this process users converted sentences going to be separate because in this process need to separate medical facts for identify diseases. Identified verbs, objects and nouns are comparing with database storing in firebase and collect only important medical facts from those sentences. Gradient Algorithm is choosing for the medical facts' dividend process.

Generate Query from Identified Facts

From those selected medical facts data should need to generate queries from these all things.

Create Flutter application to collect data about medical centers

create simple flutter application for collect data about medical centers and all of these data will be saved in firebase database table,

4. METHODOLOGY

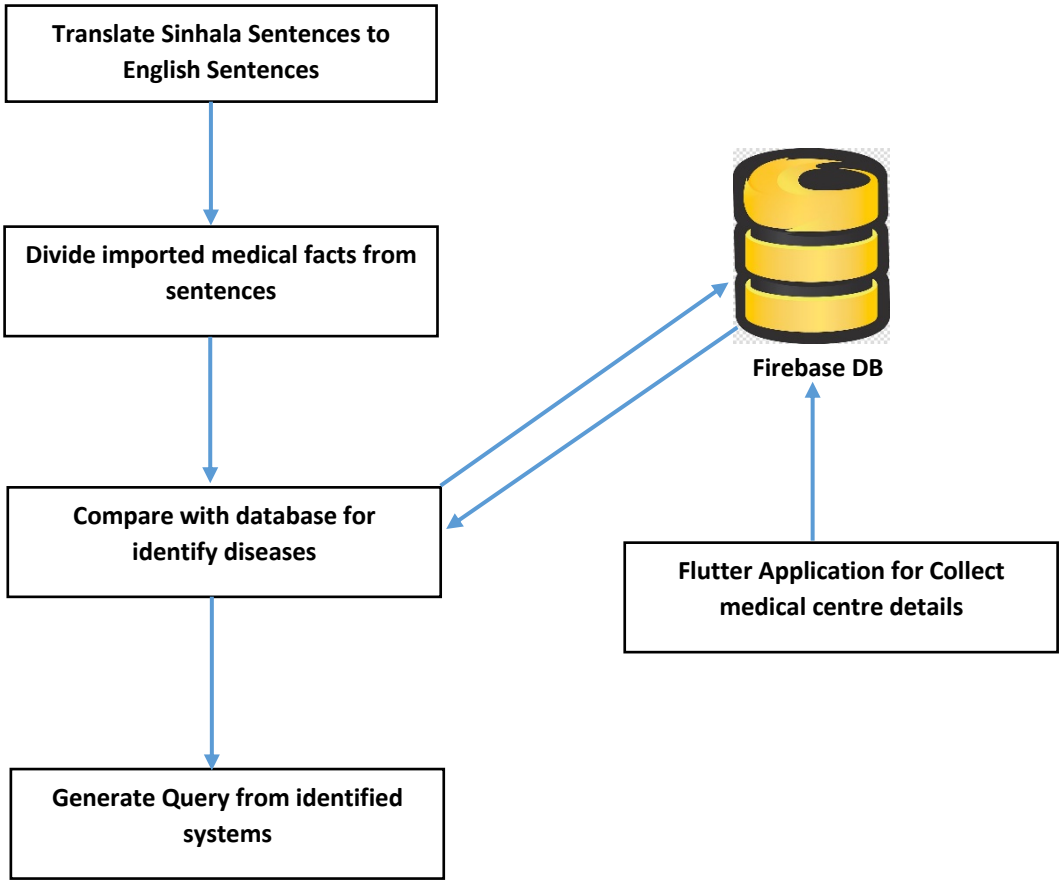

FIGURE 5 : SYSTEM OVERVIEW DIAGRAM

Create Database is an organized collection of structured information, or data, typically stored electronically in a computer system. A database is usually controlled by database management system. Database management system stores, organizes and manages a large amount of information within a single software application. In our application we have to make database with huge capacity of data. These data should retrieve fast from database. So, we decide to create our database on firebase. The main thing is firebase database easy to Connect with flutter application and it will be NoSQL database.

Natural Language Processing, usually shortened as NLP is part of artificial intelligence that works as interaction between computers and human using the natural Language. The final objective of NLP is to Read, Decipher, Understand and make sense of the human language in a manner that is valuable. Most NLP techniques rely on machine learning to derive meaning from human language. In this research NLP use for understand given English sentence.

FIGURE 6 : NATURAL LANGUAGE PROCESS

- **Structure extraction** – identifying fields and blocks of content based on tagging
- **Identify and mark sentence, phrase, and paragraph boundaries** – these markers are important when doing entity extraction and NLP since they serve as useful breaks within which analysis occurs.
- **Language identification** – will detect the human language for the entire document and for each paragraph or sentence. Language detectors are critical to determine what linguistic algorithms and dictionaries to apply to the text.

  - Open source possibilities include Google Language Detector or the Optimize Language Detector or the Chromium Compact Language Detector

Fact Identification is important to separate important words from given sentences for make answers. This application has to identify filtered English medical domains from given English sentences. Those medical domains should compare with database.

- **Tokenization** – to divide up character streams into tokens which can be used for further processing and understanding. Tokens can be words, numbers, identifiers or punctuation (depending on the use case)

  - Open source tokenizers include the Lucene analyzers and the Open NLP Tokenizer.

  - Basis Technology offers a fully featured language identification and text analytics package (called Rosette Base Linguistics) which is often a good first step to any language processing software. It contains language identification, tokenization, sentence detection, lemmatization, decompounding, and noun phrase extraction.

- Search Technologies has many of these tools available, for English and some other languages, as part of our Natural Language Processing toolkit. Our NLP tools include tokenization, acronym normalization, lemmatization (English), sentence and phrase boundaries, entity extraction (all types but not statistical), and statistical phrase extraction. These tools can be used in conjunction with the Basis Technology' solutions.

- **Acronym normalization and tagging** – acronyms can be specified as "I.B.M." or "IBM" so these should be tagged and normalized.

  - Search Technologies' token processing has this feature.

Automatically Query Generation is most important part in here. Because from these generated queries will lead to make responses. So there should have 100% correct query. The queries will have generated automatically from selected medical facts in fact identification.

Create Flutter Application, in ADAM application have to collect some information about medical centers and about doctors and their channeling times so Research Team have to implement another application for collect data. Research team is going to create it with flutter and data save in firebase database.

5. PROJECT REQUIREMENTS

Require engineering is a process of establishing

- The services that a customer requires from a system and
- The constraints under which it operates and is developed.

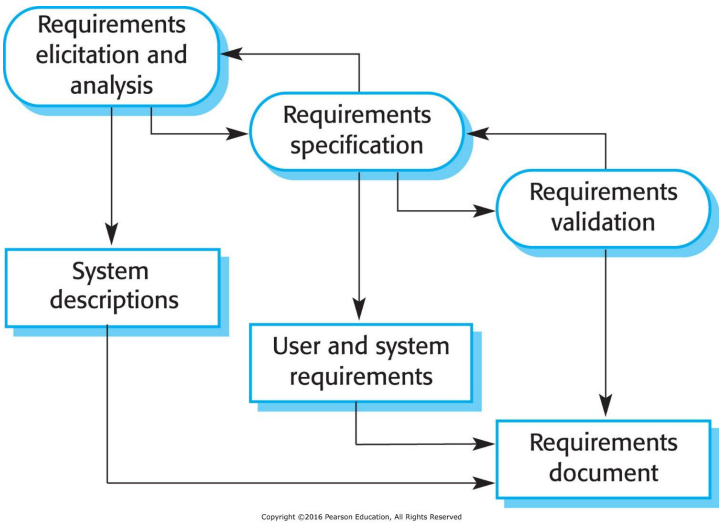

FIGURE 7 : REQUIREMENTS ENGINEERING PROCESS

Two types of Requirements Levels

1. User requirements
   - Platform to patient to e-channeling
   - Platform to dispensary/ hospital to register their system
   - Platform to doctor to register their system

2. System requirements

   A structured document setting out detailed descriptions of the system's functions, services and operational constraints

   System requirements are the configuration that a system must have in order for a hardware or software application to run smoothly and efficiently. Failure to meet these requirements can result in installation problems or performance problems.

   The former may prevent a device or application from getting installed, whereas the latter may cause a product to malfunction or perform below expectation or even to hang or crash.
   There are two types of System Requirements

   1. Functional requirements
   - ❖ Statements of services of the system should provide, how the system should react to inputs and how the system should behave in particular situations.
   - ❖ Describe functionality or system services
   - ❖ Depend on the type of software, expected users and the type of system where the software is used
   - ❖ Functional user requirements may be high-level statements of what the system should do
   - ❖ Functional system requirements should describe the system services in detail

      - Patient registration to the system
      - Patient assistant through the system
      - User interact with system using Sinhala voice command
      - Doctor appointment for the patients
      - Patients can get the information about doctors, hospitals and channeling
      - Doctor channeling through the system
      - Medical service centers registration to their mobile application

2. Non-functional requirements

- ❖ Constraints on the services or functions offered by the system often apply to the system rather than individual features or services. These define system properties and constraints
- ❖ Non-functional requirements may seem to be more critical than functional requirements. If these are not met, the system may be useless.

    - Automated conversational mobile application
    - The quality of Sinhala Voice commands and the performance like System response time, throughput, utilization, static, volumetric
    - Reliability of the system data is the most important fact for both Patients and Medical Service Centres
    - System and its whole functionality (doctor channelling, information accessibility) should be available anytime for the patients.
    - How much easier to users when interact with this system's functionalities
    - Total cost for system implementation, developments and use of the services
    - How the system localization for Sinhala users with interaction of Sinhala voice commands

Property	Measure
Speed	Processed transactions/second User/event response time Screen refresh time
Size	Mbytes Number of ROM chips
Ease of use	Training time Number of help frames
Reliability	Mean time to failure Probability of unavailability Rate of failure occurrence Availability
Robustness	Time to restart after failure Percentage of events causing failure Probability of data corruption on failure
Portability	Percentage of target dependent statements Number of target systems

FIGURE 8 : METRICS FOR SPECIFYING NON-FUNCTIONAL REQUIREMENTS

6. WORK BREAKDOWN STRUCTURE

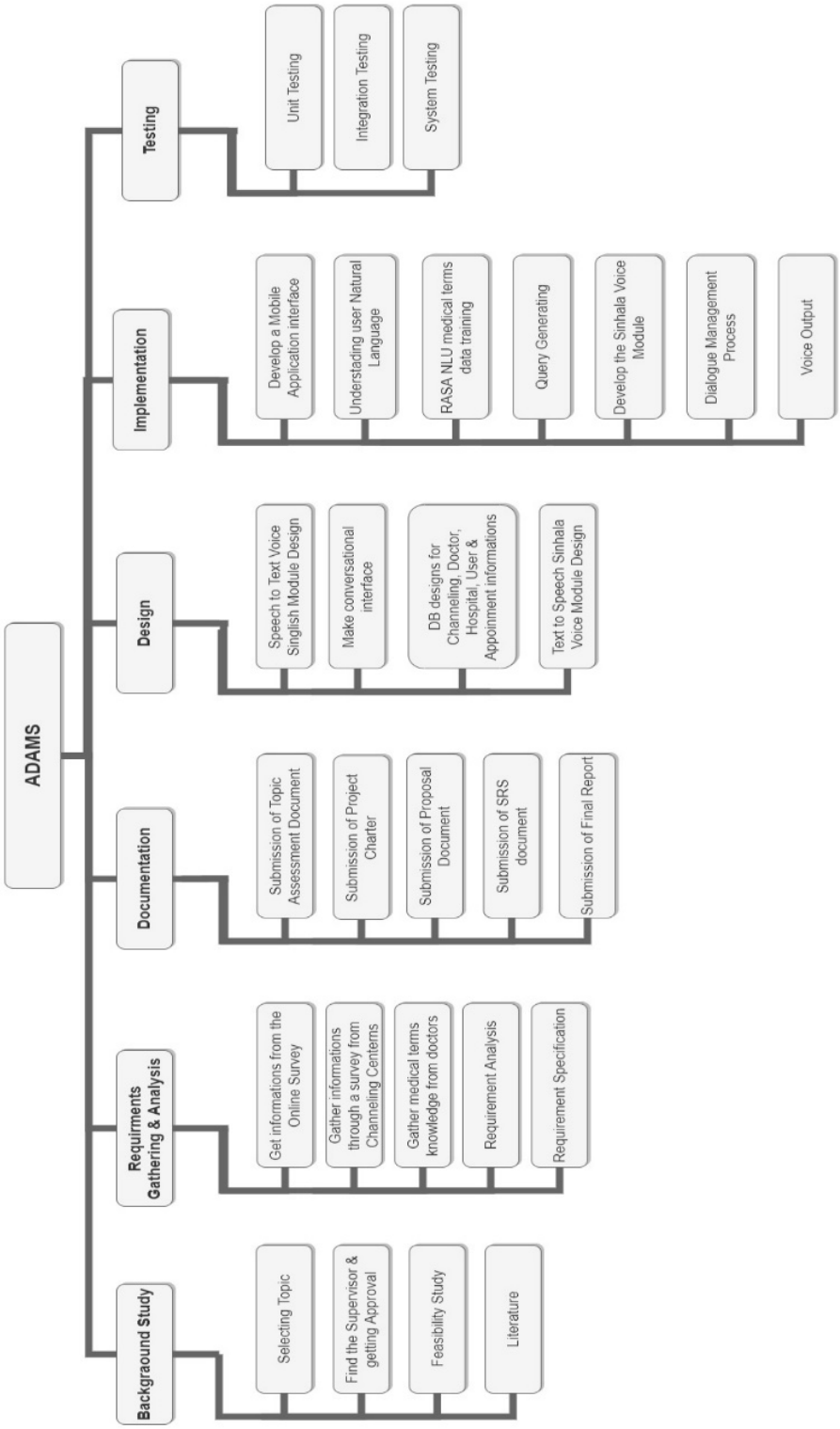

FIGURE 9 : WORK BREAKDOWN STRUCTURE

7. DESCRIPTION OF PERSONAL AND FACILITIES

- Understanding user Natural Language
- Natural Language Process
- Dialogue Management
- Connecting an NLU into Rasa framework

This project would be done by a group of four members and the research and development workload of the project is being distributed among all the members. Detailed explanation of the assigned components has been discussed in the previous sections of the document.

Member	Component
D.D.S Rajapakshe	Understanding user Natural Language
Kudawithana K.N.B.	Natural Language Process
U.L.N.P. Uswatte	Dialogue Management
Nishshanka N.A.B.D	Connecting an NLU into Rasa framework

TABLE 2 : MEMBER DETAILS WITH MAIN OBJECTIVE

8. BUDGET AND BUDGET JUSTIFICATION

This will be the budget that going to estimate in this project when implement this functionality. In Research Project is planned to use Firebase Realtime database. Because it's a mobile compatible database. Firebase has two types of database planning. First one is Spark Plan and the second one is Blaze Plan. Spark Plan is a free plan, but it can't customize. Then the best plan is Blaze Plan. Because it can customize as the developer wish.

Firebase Cloud Storage:
For the purpose of store Sinhala voice modules in cloud storage is the use of Firebase Cloud Storage. 5GB free data has already been provided for the cloud storage. Mainly in this research wants This cloud storage also provides a 5GB free data. This cloud storage also provides us a 5GB free data. Mainly store two voice modules then we want some high capacity storage. Mainly we want to store two voice modules then high capacity storage must be needed. Then 95 GB extra storage must be added.
Then total storage capacity is 100 GB. Then $2.47must be paid additionally. 30 GB free storage for data transfer and 2,100,000 operations are provided additionally.

Budget		**Cost**
1. Firebase Cloud Storage		
1.1. Storage(100GB)	$2.47	$2.47
1.2. Real Time Database	$45	$45
2. Stationaries (12 months)		
2.1.A4 sheets	$1.38	
2.2.Rough sheets	$0.55	
2.3.Binding	$1.38	
2.4.Pens & Pencils	$0.83	
2.5.Other Stuff	$1.10	$5.24
3. Communication (12 months)		
3.1.Internet service	$13.78	
3.2.Phone Cost	$16.54	$30.32
4. Printing Cost (12 months)		
4.1.Reports	$2.76	

TABLE 3 : BUDGET PLAN

9. TIME PLAN

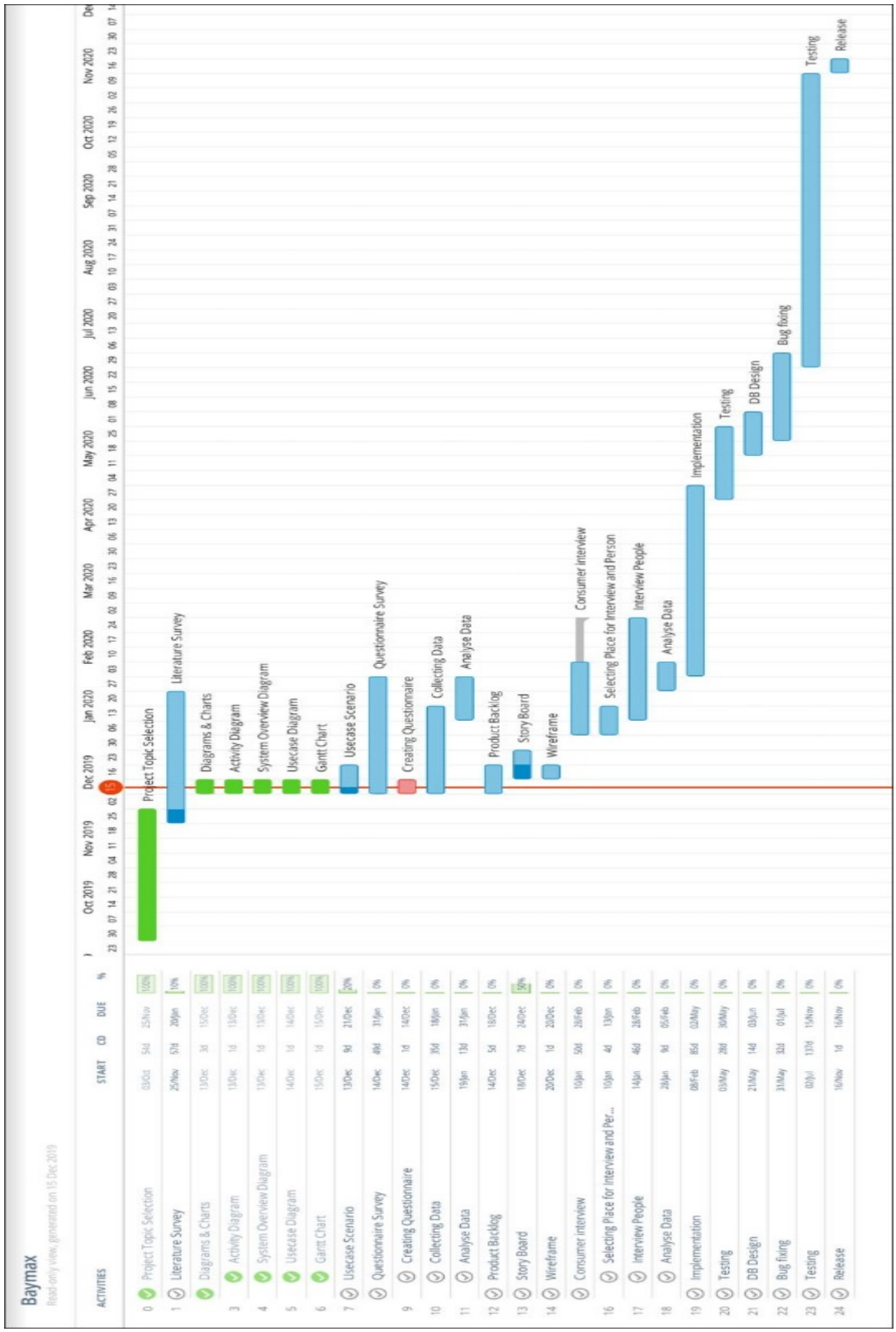

FIGURE 10 : GANTT CHART

(https://drive.google.com/open?id=1GXsOwinBr2liyusBCX8MbBiUnvq0dlom)

10 REFERENCE

[1] R. Hall´en, “http://lup.lub.lu.se/,” 02 03 2017. [Online]. Available: http://lup.lub.lu.se/luur/download?func=downloadFile&recordOId=8904399&fileOId=8904400. [Accessed 10 02 2020].

[2] A. A. Imam, “SpringerLink,” 06 December 2018. [Online]. Available: https://link.springer.com/article/10.1186/s40537-018-0156-1. [Accessed 10 02 2020].

[3] M. J. Mior, “researchgate.net,” University of Waterloo, June 2014. [Online]. Available: https://www.researchgate.net/publication/264859776_Automated_Schema_Design_for_NoSQL_Databases. [Accessed 12 February 2020].

[4] Q. Do, “https://www.researchgate.net/,” New Mexico State University, June 2014. [Online]. Available: https://www.researchgate.net/publication/273118857_Automatic_Generation_of_SQL_Queries. [Accessed 13 February 2020].

[5] B. Hettige, “http://pdfs.semanticscholar.org/,” University of Sri Jayewardenepura, 31 October 2008. [Online]. Available: https://pdfs.semanticscholar.org/aa31/76391bb4ba25c34a379f9153bcb1c5c30a6c.pdf?_ga=2.101044161.1300842029.1582446060-1805892879.1580010100. [Accessed 14 February 2020].

[6] F. H. A. M. Silva, “www.researchgate.net,” University of Colombo School of Computing, October 2008. [Online]. Available: https://www.researchgate.net/publication/266221116_Example_Based_Machine_Translation_for_English-Sinhala_Translations. [Accessed 15 February 2020].

[7] B. Hettige, “https://www.academia.edu/,” University of Moratuwa, December 2010. [Online]. Available: https://www.academia.edu/4672230/Computational_model_of_grammar_for_English_to_Sinhala_Machine_Translation. [Accessed 16 February 2020].

[8] C. Segura, “researchgate.net,” Agency for Science, Technology and Research (A*STAR), May 2018. [Online]. Available: https://www.researchgate.net/publication/327426302_Chatbol_a_chatbot_for_the_Spanish_La_Liga. [Accessed 17 February 2020].

[9] C. Liyanage, “researchgate.net,” University of Colombo, March 2012. [Online]. Available: https://www.researchgate.net/publication/235931895_A_Computational_Grammar_of_Sinhala. [Accessed 18 February 2020].

[10] B. Hettige, “https://www.researchgate.net/,” University of Sri Jayewardenepura, September 2006. [Online]. Available: https://www.researchgate.net/publication/4248469_A_Parser_for_Sinhala_Language_-_First_Step_Towards_English_to_Sinhala_Machine_Translation. [Accessed 15 February 2020].

[11] D. Khurana, “https://www.researchgate.net/,” Manav Rachna International University,India, August 2017. [Online]. Available: https://www.researchgate.net/publication/319164243_Natural_Language_Processing_State_of_The_Art_Current_Trends_and_Challenges. [Accessed 19 February 2020].

[12] D. Guo, “https://www.aclweb.org/,” Sun Yat-sen University, [Online]. Available: https://www.aclweb.org/anthology/D18-1188.pdf. [Accessed 20 February 2020].