IOT DEVICE TO DETECT ANEMIA

Project ID: 19-129

Preliminary Progressive Report – Creation of diagnosing algorithm and the backend of the system

Comprehensive Design/Analysis Project

H.S.M.H Fernando IT16136120

B.Sc. (Hons) in Information Technology Specializing in Computer Systems and Network Engineering

Department of Information Systems Engineering

Sri Lanka Institute of Information Technology Sri Lanka

May 13, 2019

IOT DEVICE TO DETECT ANEMIA

Project ID:19-129

Creation of diagnosing algorithm and the backend of the system

Preliminary Progress Review Report (Preliminary Progress Review Report submitted in partial fulfilment of the requirement for the Degree of Bachelor of Science Special Honours in Information Technology)

H.S.M.H Fernando IT16136120

Supervisor Mrs. S. Lokuliyana

B.Sc. (Hons) in Information Technology Specializing in Computer Systems and Network Engineering

Department of Information Systems Engineering

Sri Lanka Institute of Information Technology

Sri Lanka

May 13, 2019

DECLARATION OF THE CANDIDATE & SUPERVISOR

I declare that this is my own work and this Preliminary Progress Review Report does not incorporate without acknowledgement any material previously submitted for a degree or diploma in any other university or Institute of higher learning and to the best of my 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.

Author:

H.S.M.H. Fernando

IT16136120

Signature : ……

The above candidate is carrying out research for the undergraduate dissertation under my supervision.

Signature of the supervisor: …… Date: ……

TABLE OF CONTENTS

LIST OF FIGURES

1. Introduction

1.1 Purpose

The document preliminary progressive report is designed to provide a detailed document about the sub goals that must be accomplished in the proposed Anemia detection system's diagnosing algorithm and the backend, and to give an overview of the implementation of the system relating to those goals. The system portrayed in the document is and IoT device which can take physical measurement of the skin color on an index finger from a person and combine that measurement with other user information gathered via a mobile app to come up with an anemia diagnosis. This document also includes some graphs and figures that clearly describe overall idea of this research part. Interested set of audience to this document is are the supervisor, the co-supervisor, project coordinator and developers.. This document can be used as a guidance to the researchers and developers who will work on this area in the future.

1.2 Scope

The scope of this document is to provide a detailed document of research area. One of the core components in out research project is developing an algorithm which can determine whether a user has anemia or not by analyzing the inputs of the user and the implementation of the backend which will provide interfaces to the IoT device and the mobile app to carryout the work of the anemia detection system.. Statement of work describe the other related existing similar algorithms. Significance of the problem

part describes how this new algorithm is unlike other similar algorithms and the unique challenges faced in implementing this component of the project . Technical overview describes hardware and software requirements for implementing this component. Detailed design architecture shows the design architecture of the implementation and how the system component process works. Testing and

evaluation describes the how the system will be tested and evaluated against the real world anemia detection method currently used in hospitals.

1.3 Overview

Anemia is a disease that is caused due to the lack of iron in the blood which eventually reduces the blood hemoglobin level. This causes lack of oxygen in the body. The ability to detect the disease at an early stage would lead to an early cure. The main method of detecting anemia requires drawing blood from a person. This can be dangerous especially in low income countries like Sri lanka where infections can happen due to this method. Also it is preferable to avoid drawing blood from a vulnerable person such as a pregnant mother or a child.

In order to achieve this, the team will be designing a non-invasive system which has the ability to detect anemia with a high accuracy rate and giving the output to the patient so that the treatments can be taken an early stage. A device is built using a non-invasive method so that blood should not be taken and then data is sent to the server for processing and the user has to fill details in the app as well. Then the system able to analyse the data and get the results to the app.

2 . Statement of work

2.1 Overview of related previous work based on the literature survey and background information

In the invasive anemia detection methods, a device has been developed containing electronic instrumentalization, post processing software and plug and play disposable sensor which can be used to take 50µL whole blood sample to test for anaemia and provide instantaneous results[1]. Although proper testing has resulted in only 2% accuracy error , this being a noninvasive method does not solve our search

problem.Non invasive devices like project that is done by R. Bhattacharyya which is using RFID the ability to test for anaemia[2] exists but the accuracy is not comparable to the traditional methods. This is why we are proposing a system where a device is present to take measurement from user and then that data is processed using image processing inside the device. Then that data is used together with user inputs through a questionnaire in the mobile app and all that data together is put through an algorithm to produce an outcoming determining whether the user has anemia or not. By doing research we can reasonably determine that a machine learning algorithm would be best suited for this type of work[3]. So I have chosen supervised machine learning algorithm type which is a classification algorithm. In implementing this component of the project I have chosen Machine Learning as a Service(MLaaS) and Backend as a Service(BaaS) as they provide many benefits[4] of traditional methods including costs saving, high scalability and reduced workload on the team.

2.2 Identification and significance of the problem

When developing the anemia detection system, my component is focused on the algorithm which will analyse the data received from the device and the mobile app to come up with the diagnosis and the implementation of the backend which will provide interfaces to the device, app as well as other necessary services. Although there exists Machine learning algorithms and other techniques(such as Neural Networks), most if them use only the data taken by a device to produce the results. But in our project, we use multiple data inputs from the questionnaire in the mobile app in addition to the data gathered by the device to produce the result. This is unlike any other solution in the industry that is known. The data set used to train the prediction model produced by the machine learning model needs to be changed and adjusted in such a way that the best model possible is generated. As the data from image processing done at the device and the data from the mobile app is rather large in size and complexity it presents a significant challenge in obtaining an acceptable result at the completion of this component. With the scope and complexity of the input data there is also the challenge of adjusting the necessary thresholds and the parameters of the algorithm to minimize false positives and false negatives.

There also needs to a backend to handle all the data inputs and outputs from the device and the app and store them in an appropriate manner. The backend must also authenticate the device and user and must also provide management interfaces for management purposes. The challenge is there to come up with a solution that can handle the different formats of data to be exchanged with the MLaaS API and the mobile app API. The backend services must be available at all times and must be able to handle dynamic amounts of traffic.

2.3 Technical Overview

Here two solutions will be used and and after testing and taking into consideration the cost, scalability and manageability of each solution , the best one will be selected for the final implementation.

One solution will use Microsoft Azure Machine Learning service as the MLaaS and the other will use Amazon Sagemaker MLaaS. Parse server which is an open source BaaS framework will be hosted on Azure and AWS cloud respectively. Firebase and MongoDB will be sued for data storage.

To carry out initial development before moving to the cloud and to test out changes before they are pushed to the servers in the cloud, a local computer will be used with Virtual machines containing the local versions of the servers. The specifications of this computer are given below.

o Memory -8GB

o CPU- Core i5-2400 (4 cores/4 threads)

o Disk-500GB (SATA 3)

o Network-one Intel 82545EM Gigabit Ethernet NIC

3 . Research Methodology

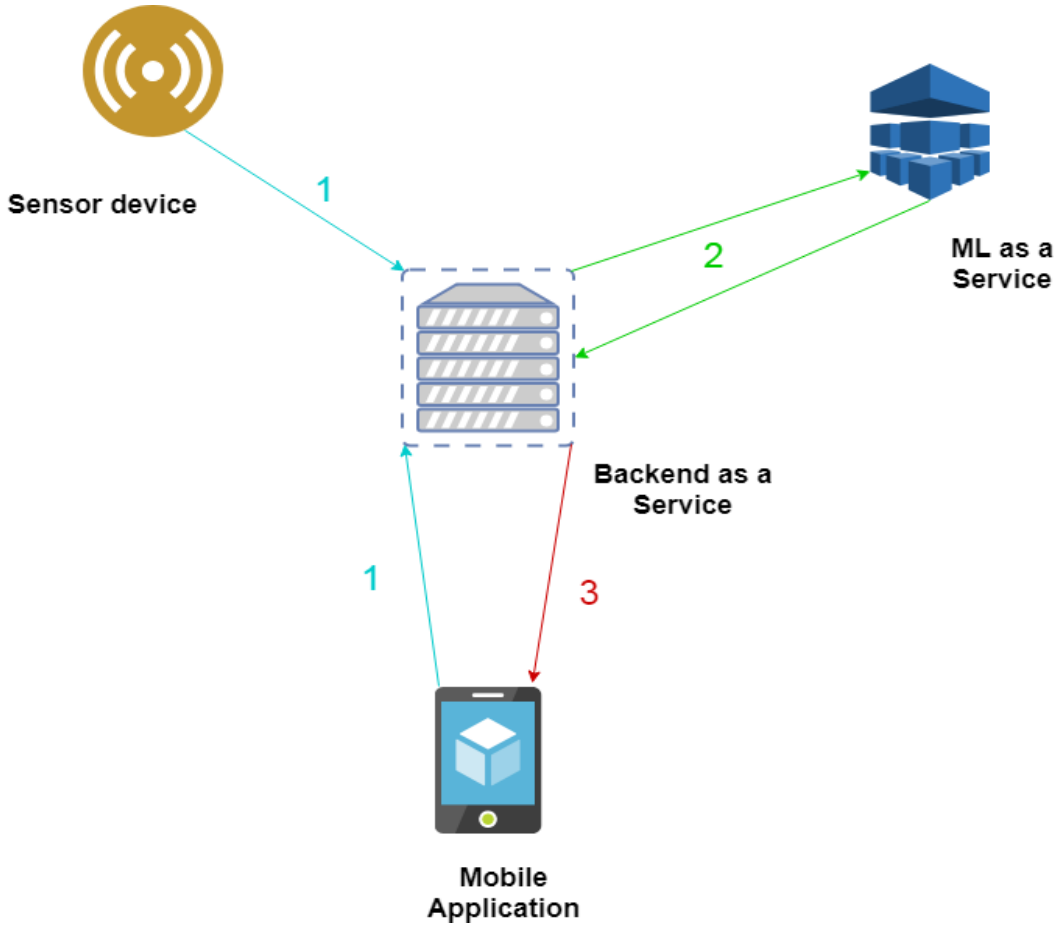

Figure 1

Here we have the server running in the BaaS and it is interfaced with the device, the mobile app and the MLaaS APIs. The main procedure of the design is as follows.

1. A. The device sends authentication data. Then the device is authenticated and the data outputted by the image processing in the device is transferred to the server.

   B. A. The mobile app sends authentication data. Then the app user is authenticated and the data gathered by the app via a questionnaire is transferred to the server .

2. Then the server will transform that data into the appropriate data set and that data set is sent to the MLaaS API to be consumed by the prediction model.

Then the prediction model will analyse the data and generate a result which will be transferred back to the server.

3. The server will store this result and the necessary data of the associated auser and send a message to the mobile app API which will then display the result to the user.
4. There will also be a management interface to manage the system.

Here a management interface will also be implemented in the server to serve the management purposes of the system. The management interface should be able to sort through the data stored and generate reports which contains statics about the system and its accuracy.

To produce a data set, data will be collected through the device and the app for a suitable sample of users who are a mixture of anemia patients and healthy persons. Then that data is used to create training sets for each iteration of the machine learning algorithm.In terms of training the machine learning algorithm a training data set will be prepared each time and 70%-80% of the set will be put through the learning algorithm. The resulting prediction model will be tested against the remaining 20-30% of the data set. Changes will be made to the training data set ,thresholds of the machine learning algorithm and parameters o the machine learning algorithm until a prediction model with satisfactory accuracy is produced.

4. Test data & analysis

As mentioned before, a portion of the training data set will be reserved to be tested for the accuracy of the result of each prediction model. From that comparison we can get a good idea of the accuracy of the model. As for final testing the whole system will be compared against real testing methods used in hospitals and s statically analysis will be performed to determine the accuracy of the system compared to traditional methods such as blood sampling.

5. Anticipated benefits

- A highly accurate non invasive anemia detection system
- Valuable data store for future research purposes
- Lower the costs incurred by doing traditional testing over time
- Valuable research paving the way for similar type research and innovation
- Reports generated can be analysed to determine trends among users

6. Project plan or schedule

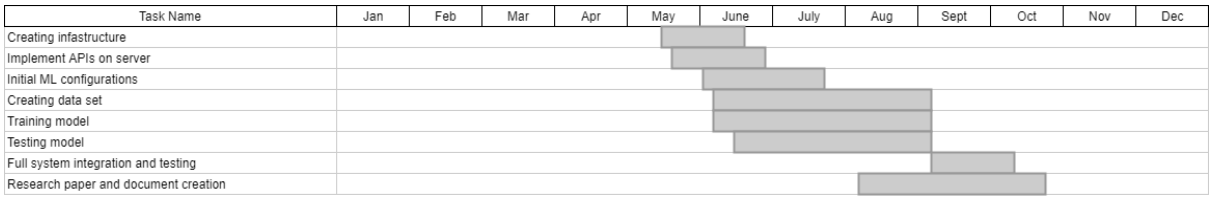

Figure 2

7. Research constraints

The data transfers between server and the device and the app must be taking as less bandwidth as possible.

The amount of simultaneous detections operations must be limited to save costs and other complexities in doing the research.

The number of people to be sampled for testing and evaluation as well as the development of the prediction model is limited as the resources available to do the sampling process is limited.

8. Specified deliverables

The primary outcome of this research project is an efficient and effective solution to detecting anemia early. Our proposed system of this research project is aimed to develop a system which can be used to detect anemia by analysing various user inputs specified. The system will determine the users' current intelligent level by testing environment.. As for the component specified in this document, a decision-making model is expected to be implemented in order to determine whether a user is suffering anemia or not and then to manage and store that user data in a useful manner.

9. References

[1] J. C. C.-F. R.-V. L. M.-C. Jaime Punter-Villagrasa, "Toward an Anaemia Early Detection Device Based on 50-µL Whole Blood Sample," IEEE, Madrid, 2015.

[2] E. A. I. E. S. S. C. S. B. T. a. M. W. R. Bhattacharyya, "Towards low-cost, wireless blood anomaly sensing:," IEEE, San Diego, 2015.

[3] M. P. Meherwar Fatima, "Survey of Machine Learning Algorithms for Disease Diagnostic," Scientific Research Publishing Inc, Multan, Pakistan , 2017.

[4] B. Schoepfle, "How "as a Service" (aaS) Model Works & How Companies Benefit from It", iT1 Source, 2019.