IoT Device to Detect Anemia: A Non-Invasive Approach with Multiple Inputs

Akmal Hafeel
Department of CSE
SLIIT
Malabe, Sri Lanka
akmalhhafeel.bc@gmail.com

M. Pravienth
Department of CSE
SLIIT
Malabe, Sri Lanka
mahespravienth123@gmail.com

N. Kayanthan
Department of CSE
SLIIT
Malabe, Sri Lanka
kayankayanthan@gmail.com

H.S.M.H. Fernando
Department of CSE
SLIIT
Malabe, Sri Lanka
mfernando3917@gmail.com

Shashika Lokuliyana
Department of CSE
SLIIT
Malabe, Sri Lanka
shashika.l@sliit.lk

Dr. Anuradha Jayakody
Department of CSE
SLIIT
Malabe, Sri Lanka
anuradha.j@sliit.lk

***Abstract*— Detection of diseases is a key aspect when considering the world today. Diseases if not cured early it gets worse day by day. Therefore, early detection of the disease is a key aspect that needs to be considered. In this project the aim of the team is to develop a device to detect Anemia which is a disease caused due to the lack of Fe3+ ions in the body. This is a disease that would even lead to organ failure eventually leading to a heart attack or death. Anemia is causing the red intensity of blood to reduce. Therefore, the team is using that factor to detect the disease. The team have built this device using a key symptom where the intensity of red in blood is measured and the output is sent to a central server, whilst the server also gets input from a questionnaire that is built in app. Using these data, the server will decide whether the user is suffering from anemia or not.**

Keywords - anemia, image-processing, Keras, machine learning, IoT

I. INTRODUCTION

Anemia, a disease that is mainly caused due to the reduction of iron in blood, which is a common blood disorder. If someone is carrying this disease, then the production of red blood cells in the body is low, which causes the protein known as hemoglobin to reduce. Reduction of hemoglobin would cause a major issue in the body, where it would reduce the amount of oxygen that would be carried in the blood would reduce. The reduction of this is a major issue because if the required amount of oxygen is not transferred to the organs then the organs would start to malfunction and sometimes organ failure could even be a possible

outcome. This is like one's body is starting to suffocate from within the body due to the lack of oxygen. This disease can be identified at the initial stage when a full blood count is taken if the amount of red blood cells is less than 12 grams per deciliter for women and 15 grams per deciliter the doctor would direct the affected69 for further testing because normally this reduction is caused due to anemic conditions in the body. But to narrow down into the root cause the medical officer would direct those patients for further testing. Some of the identification symptoms that can be seen in an anemic patient is that the patient would be having trouble in breathing. This is mainly caused due to the lack of red blood cells to carry oxygen, so the patient would develop heavily breathing conditions [1].

There are mainly 3 causes of Anemia. The first is blood loss. When a person suffers from greater blood loss then a large amount of red blood cells is lost. During that time the production is not enough to cater the amount of red blood cells the body needs. The probability of women getting anemia is high due to the blood loss during the menstrual cycle. Heavy menstrual cycle periods would cause larger blood loss making women more likely to carry this disease. Other types of blood loss could be caused due to the amount of blood that is lost during accidents and external bleeding due to surgery or due internal bleeding that is very hard to identify. The second cause of anemia is due to the reduction in the production of red blood cells. This condition can be caused by an iron lacking diet, medical condition such as cancer, AIDS or women who are pregnant also would be also anemic due lower production of red blood cells, children also can be anemic due to genetic disorders of not producing enough red blood cells. The last cause of anemia is the increased amount of red blood cells destruction by the spleen. This can be caused due to certain diseases that would cause the body to destroy more red blood cells from the body. The second reason for this to happen would be an enlarged spleen that is causing the destruction of red blood cells that would cause anemic conditions in the patients [2]. Prevention is better than cure, but early detection can get the patient suffering to the best cure. Therefore, identifying anemia at the initial stage and taking necessary treatments would make the condition reduce and stop from reaching severe conditions.

II. LITERATURE REVIEW

Due to the increment in different types of diseases there is a need for better technologies for the identification of these diseases which the patient is diagnosed from. There was a research conducted by Jamie Punter and five other members. In this research they obtain 50µL whole blood sample to test for anemia. The main aim of the project is to get a portable device so that it can be used for instantaneous detection of anemia. The developed device contains electronic instrumentalization, post processing software and plug and play disposable sensor. The disposable sensor is based on a three-gold electrode commercial sensor which is of low cost and 50µL of blood is required for the device to use in this test. In order to achieve a success with this device they had used 48 blood samples for testing of this device. These samples

were collected from different clinics and hospitals for the task. Blood samples were distributed in two main groups as one for system calibration and the other 38 samples for system validation. The calibration of the device was done using a complete EIS experiment in this project to get accuracy in the detection of anemia, defining the working range of hematocrit detection. So, in this project the specialty they had used an instant impedance detection in order to make sure they achieve accuracy, sensitivity and co-efficient of variation in this project. After the proper testing has been carried out only 2% accuracy error had been seen [3]. But the main difference between the project that is being developed, and the above mentioned project is that the research team will be also developing a portable device, but this new project is non-invasive where the team will not be getting any blood samples to carry on any tests but rather make sure that the required input is taken from the symptoms. For this a device is built with the necessary features with image processing integrated into it. The next most important difference between the device that is developed, is that the device would be containing a connection to server which in this research is residing in the cloud which the team is connecting to send the relevant data and those would be connected to the mobile device of the patient. The patient should also answer a questionnaire that is designed in the app and the accuracy of the system would also increase because many symptoms would be analyzed and mainly using the device, the team would be taking a main input which to determine whether the patient is diagnosed with this disease or not. The system is going to detect the whether the blood flow in the patient is low by using image processing and relevant technologies. Using these the team can detect the blood flow and determine the severity of anemia in the body. This would be the idlest method to detect anemia in a non-inversive methodology.

The next project that is done by R. Bhattacharyya which is using RFID the ability to test for anemia. In their project which is titled, “Towards low-cost, wireless blood anomaly sensing: An RFID-based anemia detection sensor” they use a method there they would be able to send a certain frequency to detect if there are ample amount of red blood cells in a given quantity of blood inside the body. Therefore, in order to detect that they have been using a Infra-red frequency and send it back and retrieve the relevant data and analyses whether there is a difference between the normal person’s blood and an anemia infected patient. In an anemia infected patient anyways, there will be lower red blood cell count. Therefore, in their project they have used a sensor which is capable of reliably differencing between blood having 20, 30, 40 and 50% red blood cells concentration by volume. In the RFID equipment that is used by this team has the ability of allowing for automated screening of blood specimens at large scale is also available. Therefore, according to this project the Red Blood count is taken by volume and is checked for anemia [4]. Image processing is also used in this project. Image processing is used in order to detect many kinds of changes, so in anemia image processing can also be used to detect the exact symptoms and those changes. In the project the system would be using machine learning algorithm as well. There are plenty of algorithms for implementing machine learning algorithm. In order to achieve this, the machine learning algorithm type used is supervised learning. There are projects that are done using the machine learning concept to detect diseases. For example, the project done by Meherwar Fathima on the topic, “Survey of Machine Learning Algorithms for Disease Diagnostic” [6] has used this method to detect diseases in plants. This concept is used in order to detect the diseases in users of the system.

III. RESEARCH METHODOLOGY

The main procedures that is involved in the designing of this project is to build a simple device that has accuracy and give the user the ability to trust the results of the device. In order to achieve the above-mentioned objectives, the team have designed the project on 4 major steps.

- Building the hardware device
- Configuring Image processing
- Configuring connection between server and device
- Configuring the server and creating necessary channels to intake data from app
- Building the app
- Designing an algorithm in the server to process the data using machine learning algorithms

So, the above-mentioned steps are the key steps in building up this device and establishing connectivity in order to make the users able detect whether they are diagnosed with anemia or not

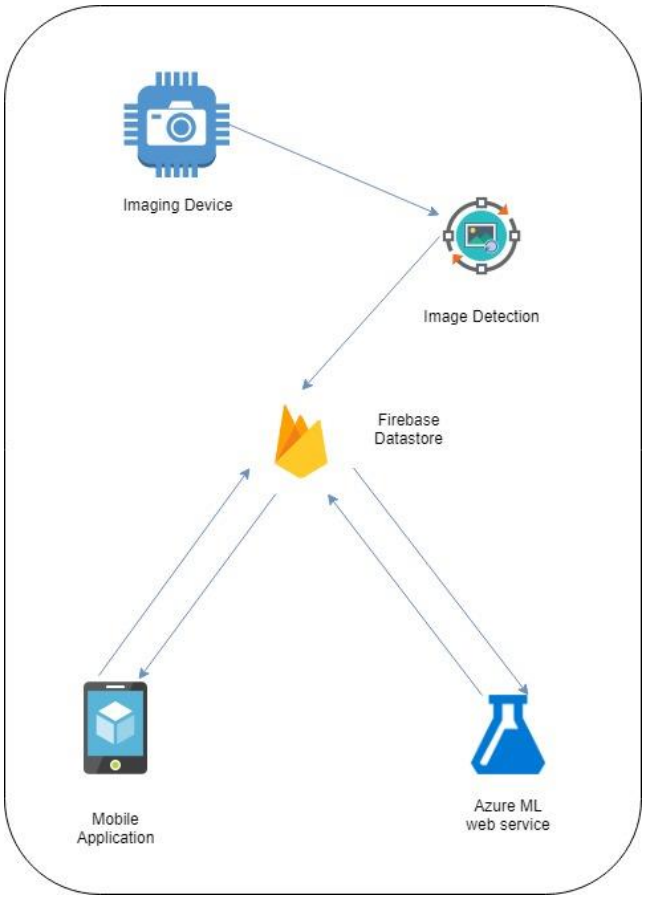

Figure 01: System Architecture

The system design architecture which is shown in Figure 01 describes the communication structure that is implemented in the project design. Initially the system is designed to get inputs from the device as well as the app and the design will be giving outputs to the app. The input that is taken from the device is the image of the fingertip. The image is then transferred to the server and the image is processed in the server. The server is subjected to image processing in the server. The server then transfers the output or the result that is obtained out of the image processing to the server where the machine learning algorithm is configured. The input which is taken via the app, through the questionnaire that is designed will also be transferred to the server. The machine learning algorithm is designed so that the inputs from the device and the app does learn the output of both anemia patient and the healthy patient so that when the device is in use the device would be configured accordingly so that an accurate output can be given. The machine learning algorithm that is configured needs to be trained so using some deep learning concepts so that an accurate result can be published. After the inputs are processed in the algorithms that is designed the output is sent to the app which is connected through an API. This would then be displayed in the app, so that the user is eligible of reading the output that is generated. The output is also recorded in the server so that the user can register an account and even when the device is broken the data can be retrieved for future use.

A. *Hardware Device Design*

Initially the device is built so that the input can be sent from the device to the servers. Therefore, the team have decided in using a raspberry pi in order to configure the device. The raspberry pi would be connected to the camera module that will be used. The camera module would be designed in manner attached to a clipping device, where the device would be clipped on to the hand of the user and a picture of the finger would be taken. In order to take that the camera should be configured so that it is able to take the pictures in a more focused manner so that the level of identification is high. After the pictures are taken then they should be sent to the raspberry pi where the team is implementing image processing. So, when designing the device, the clipping should be done so that the finger is tightly held and then released upon the release during a specific period another snapshot of the finger should be taken. In this device a strong light would also be fixed so that it passes through the finger where it would make the blood visible for the camera and when the finger is tightened then the blood moves back and that also can be noticed. Depending on that the images would be loaded in image database for image processing. The pictures would be eventually configured.

B. *Image Processing*

The image which is taken from the designed device is transferred to the server and is subjected to image processing. In order to configure image processing in the server the team is using the Keras model. The model is a design of deep learning. The model shown in Fig 02 describes how the keras model functions.

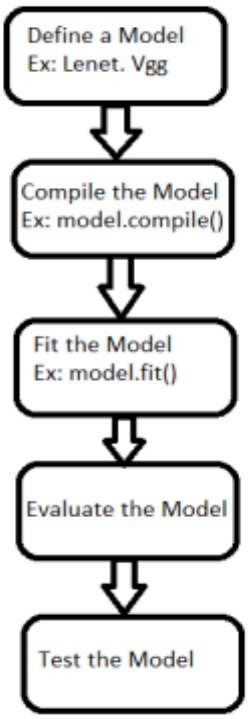

Figure 02: Keras Model

According to the above model the image processing algorithm would be implemented. The data images of the anemia patients and the healthy patients are fed into this machine algorithm for it to learn the sequence of design. Then the algorithm is run against a test data sample to increase the accuracy. The model would generate a high accuracy rate compared to other models as it is a deep learning algorithm.

Below are the samples of the images taken of healthy person and anemia diagnosed patient which is fed into the system to train. In this stage the color of the image is tracked. The heathy individual should have dense red and the sick patient has a yellowish color in the blood the change is tracked. The model is trained accordingly based on the color.

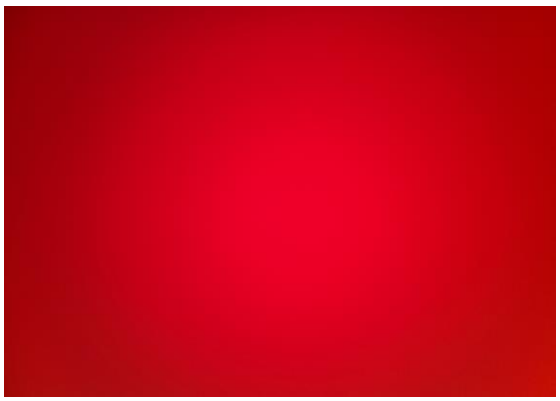

Figure 03: Healthy individual sample

The sample for the healthy patient can be seen in a dense red color shown in Fig 03. The color of the anemia diagnosed patient's sample is shown below.

Figure 04: Anemia individual's sample

The color change which can be seen in the above image can be tracked and the model is trained to track the difference. The difference is then recognized by the model and the main aim of using the deep learning machine learning algorithm is to detect the difference to identify anemia patients.

C. *Server Connectivity*

The connection should be established between the server device and app should be established. The cloud server is used to host the image processing algorithm and the image captured in the device is transferred through RSA authentication method to the server. The server then transfers the output from the image processing algorithm to the machine learning component of the server. The server and the app are also connected using API's. The app transfers the data from after the questionnaire is answered to the server and the server takes the input from the app as well as the output from the image processing model and is trained to decide whether the user of the device is diagnosed with anemia or not. Firebase functions [5] are used to trigger actions in response to updates to the database.

D. *Design of the Mobile Application*

The mobile application is based on java IntelliJ platform. The mobile application is named as "Doc Detector". It has main components such as Doctor Login, Admin Login, User Login, Patient details, Doctor details, questionnaire and finalized results page. The doctor is able to register patients for each doctor zone. Patients results can update in every month or once a week. It depends on their disease levels. Patients are able to update their profile in Doctor's mobile application. Another side of the mobile application is Admin's page. He/ She can update or edit doctors and patient details and able to edit their account. The main component is questionnaire and its shows the fifteen questions about symptoms of Anemia disease. It builds the half results of the disease and rest of the result is calculated by the device. It calculates and sends it to the mobile application for the final calculations. This will help to finalize the results to calculate the finalized results. The design of the User Interface of the questionnaire is shown below.

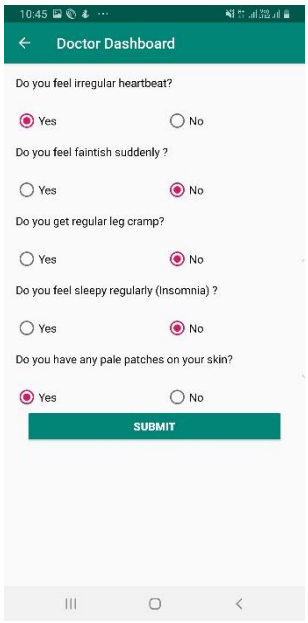

Figure 05: UI of the questionnaire

E. *Algorithms to Process Data*

Based on the inputs given by the app and the device, the device that will be designed so that it directs the input to the server and the app will also be configured with the server in order to fulfil the required information that would be committed as input for the server. The server then processes the information that is taken from both inputs. In order to achieve this, the team is implementing supervised machine learning. In this step the team is using a methodology where a set of training data taken from the output of the image processing algorithm and the questionnaire output is used to achieve this. This training data for the algorithm to run should be taken as a faithful information to be collected and run in these algorithms. In these algorithms the server would play a major role so that the inputs taken would be from data that has already been collected from patients diagnosed with anemia as well as a test from a normal healthy person, so that the processing machine learning algorithm [6] can be trained to each data set, which would expand the chances of getting more accurate results is high. The specific machine learning algorithm to be used for the system was determined to be a 'two-class decision jungle' algorithm [7]. Unlike conventional decision trees, a Directed Acyclic Graph (DAG) in decision jungle algorithms allows multiple paths from root to each leaf. This makes the algorithm to be more suited for the sample size of training data are collected in this project. The following figure shows how a DAG is formed with each iteration of the algorithm.

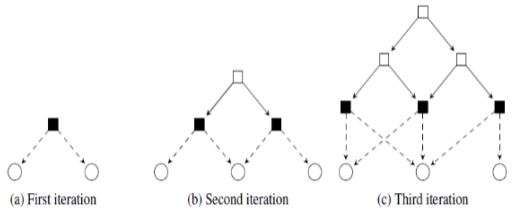

Figure 06: Iterations of a decision jungle [8]

The algorithm used for processing data is a two-class decision jungle machine learning algorithm by Microsoft. In this algorithm each rooted decision DAG in a jungle is trained independently [8]. Training is done by growing a DAG one level at a time where at each level, the algorithm jointly learns the features and branching structure of the nodes. This is done by minimizing an objective function defined over the predictions made by the child nodes emanating from the nodes whose split features are being learned. Consider the set of nodes at two consecutive levels of the decision DAG where the set of parent nodes is N_p and the set of child nodes is N_c. Let Θ_i denote the parameters of the split feature function f for parent node $i \in N_p$, and S_i denote the set of labelled training instances (x; y) that reach node i. Given i and S_i, computations can be done such that the set of instances from node i that travel through its left and right branches as $S^{L_i}(\Theta_i) = \{(x; y) \in S_i \mid f(\Theta_i; x) \leq 0\}$ and $S^{R_i}(\Theta_i) = S_i / S^{L_i}(\Theta_i)$. $l_i \in N_c$ is used to denote the current assignment of the left outwards edge from parent node $i \in N_p$ to a child node, and similarly $r_i \in N_c$ for the right outward edge. Then, the set of instances that reach any child node $j \in N_c$ is [8]:

$$S_j(\{\theta_i\},\{l_i\},\{r_i\}) = \left[\bigcup_{i \in N_p \text{ s.t. } l_i=j} S_i^{\mathrm{L}}(\theta_i)\right] \cup \left[\bigcup_{i \in N_p \text{ s.t. } r_i=j} S_i^{\mathrm{R}}(\theta_i)\right]$$

The objective function E associated with the current level of the DAG is a function of $\{S_i\}_{j \in Nc}$. The problem of learning the parameters of the decision DAG as a joint minimization of the objective over the split parameters $\{\Theta_i\}$ and the child assignments $\{l_i\}$, $\{r_i\}$ can be formulated. Thus, the task of learning the current level of a DAG can be written as [8]:

$$\min_{\{\theta_i\},\{l_i\},\{r_i\}} E(\{\theta_i\},\{l_i\},\{r_i\})$$

The figure below shows a DAG visualization [9] with the different colors indicating likely classes at each node and the saturation indicating purity.

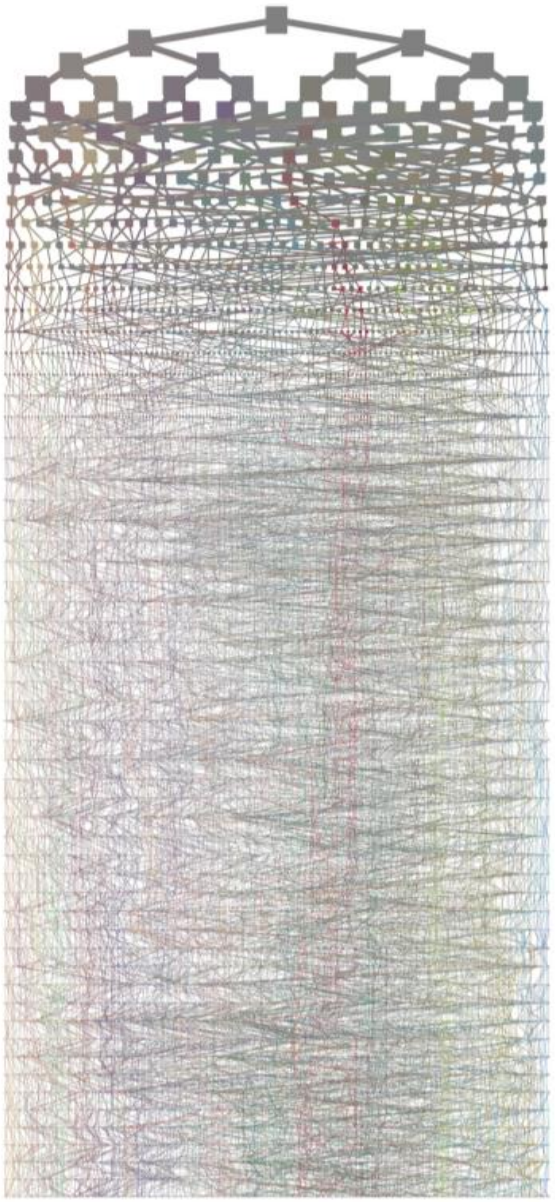

Figure 07: DAG visualization [9]

IV. Results and Discussions

After the implementation of the device the research team was able to test it on a certain number of individuals who were healthy as well as who were diagnosed with anemia. The results were initially showing an accuracy rate of as shown below. But with continuous training that the device was subjected to it was able to show an accuracy rate which is more than 80%. The device has then been configured with an external design so that it can be marketed at a later stage. When configuring the above-mentioned models there were certain problems which would be faced when configuring the image processing as the data samples which has been collected for the implementation of the demo is not exceeding more than 150. The amount of data sets that should be collected in order to train the system is manually done as the data set is not available therefore if a large sample of dataset can be collected then it could be used to train the system more accurately that would make the system eligible

to attain a greater result with a better accuracy. But currently due to the low amount of data that is available a low amount of accuracy is arrived at. The ability to make the device portable and the usage of non-invasive method is an attractive feature that leads more users to suggest the device. This reduces the blood loss in anemia patient facilitating the users the device fearlessly. The data processing algorithm to train the predictive model which produces the results was set up in Microsoft Azure Machine Learning Studio as shown in the figure below.

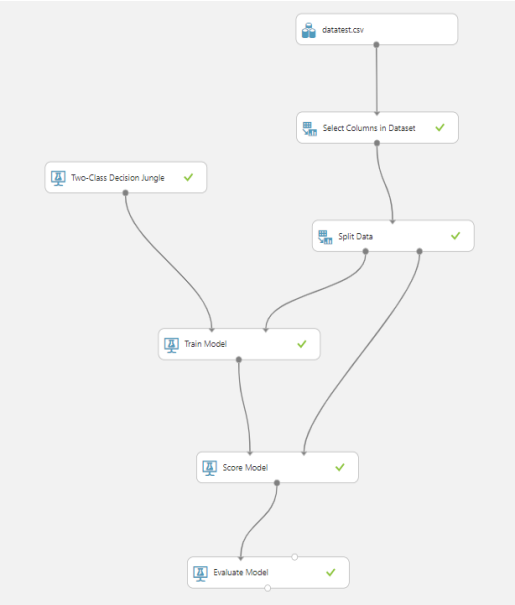

Figure 08: Predictive model structure

The dataset is split into two portions in 80:20 ratio and the first portion are used to train the model while the latter portion is reserved to test the produced predictive model in order to measure its accuracy.

The result of the evaluation is shown in the figure below with the predictive model achieving an accuracy of 82.6%.

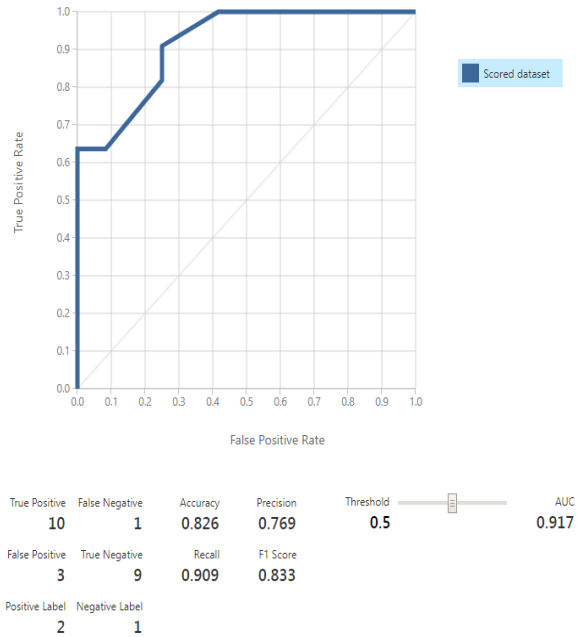

Figure 09: Predictive model evaluation

V. Conclusion and Future Works

In the research project anemia can be identified in a non-invasive method so that the blood loss can be minimized for an anemia patient. The implementation of questionnaire which collects more symptoms related to anemia and not only depending on the device. But it also takes multiple inputs from the device which makes the accuracy rate higher than only getting one symptom out of the device.

As an improvement for the device this can be designed to read the amount of hemoglobin level in the blood. If the amount of hemoglobin level in the blood can be detected, then some major diseases can be detected easily.

Acknowledgment

This work was supported by Sri Lanka Institute of Information Technology

References

[1] D. Thompson Jr., "Are you anemic," Everyday Health, 01 September 2014. [Online]. Available: https://www.everydayhealth.com/anaemia/anaemia-basics.aspx. [Accessed 28 July 2019].

[2] J. Cafassoa, R. Nall, "Iron Deficiency Anaemia," Healthline, 17 July 2017. [Online]. Available: https://www.healthline.com/health/iron-deficiency-anaemia. [Accessed 28 July 2019].

[3] J. Punter-Villagrasa et al., "Toward an Anemia Early Detection Device Based on 50-µL Whole Blood Sample," IEEE, Madrid, 2015.

[4] R. Bhattacharyya et al., "Towards low-cost, wireless blood anomaly sensing:," IEEE, San Diego, 2015.

[5] Firebase, "Firebase, Cloud Functions for Firebase |," Firebase, 2019. [Online]. Available:https://firebase.google.com/docs/functions. [Accessed 01 August 2019].

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

[7] J. Shotton et al., "Decision jungles: Compact and rich models for classification," Microsoft Research , US, 2013.

[8] T. Pohlen, "Decision Jungles," Computer Vision and Pattern Recognition, 2014.

[9] J. Shotton et al., "Decision Jungles: Compact and Rich Models for Classification Supplementary Material," Microsoft Research, Cambridge, UK, 2013.