

MOM

Pregnancy Complication Prediction Based on Machine Learning

Aleena A Thomas¹, Anas Bin Azeez¹, Fathima Sona¹, Jithin MS¹, Jinu Thomas²
Computer Science and Engineering Department, Saintgits College of Engineering
Kottukulam Hills, Pathamuttom, Kottayam, India

Abstract— Pregnancy also known as Gestation is a period of time where a woman undergoes many physiological changes. Blood sugar, pressure and cardiac output are all required and regularly check-up are needed. But sometimes abnormalities occur which induces to severe conditions which may lead to complications during pregnancy. Pregnant women must be taken more care and thus be protected from the complication that arises during pregnancy period. This project provides a self-assessment of the pregnancy providing regular monitoring prediction about the pregnant lady and the fetus. Classifications of the several algorithms are successfully implemented in this project which focuses on the machine learning.

With the improved technology this project benefits to people in the urban areas as well as to the people who are not able to visit the hospitals for routine checkups. The vital data readings are measured using Arduino. These measured data are stored in the cloud which contains machine learned data's. These machine learned data's are trained and arranged according to the high priority data's which are analysed by a doctor. This project highlights the effective performance accurately which can predict the risk levels during pregnancy from the collected, standardized and transformed data very efficiently. Thus, with this project the chance of risks, complication can be avoided for the upcoming mothers.

Keywords—Gestation, Health monitoring, Machine Learning, Risk Prediction, Arduino

I. INTRODUCTION

In the urban areas mostly 60% of the women do not take the medicines at the proper time during gestation. Pregnancy is the time where a woman undergoes with physical and psychological changes. The women must be taken care during the gestation period as some severe conditions may arise which may lead to the death of the mother or the fetus. Machine learning is a technology which turned to be very useful for the women undergoing the pregnancy. Machine learning is the subfield of computer science that gives computers the ability to learn without being explicitly programmed. Evolved from the study of pattern recognition and computational learning theory in artificial intelligence, machine learning explores the study and construction of algorithms that can learn from and make predictions on data. Machine learning is closely related to computational statistics, which also focuses in prediction-making through the use of computers. Machine learning tasks are typically classified into two broad categories:

A. Supervised learning: Supervised learning is the machine learning task of inferring a function from labelled training data. The training data consist of a set of training examples. In supervised learning, each example is a pair consisting of an input object and a desired output value. A supervised learning algorithm analyses the training data and produces an inferred

function, which can be used for mapping new examples. Supervised learning includes Classification techniques.

B. Unsupervised learning: Unsupervised machine learning is the machine learning task of inferring a function to describe hidden structure from unlabelled data. Since the examples given to the learner are unlabelled, there is no error or reward signal to evaluate a potential solution. Unsupervised learning includes Clustering techniques.

Pregnancy induces changes in women body both externally and internally. Some changes are so dangerous resulting in losing maternal and fetal lives. According to a few gynaecologists it is defined every trimester to be around 14 weeks, adding up to around 42 weeks of pregnancy.

The most common reasons for complications and deaths during period of pregnancy include maternal sepsis, maternal bleeding, abortion complications, and high blood pressure in pregnancy. gestational diabetes and weight gain during period of pregnancy. Among these complications abnormalities in Blood Pressure (BP), Blood Glucose Levels and Weight Gain are considered the most dangerous and preventable complications.

There is increased possibility of restricted fetal growth and birth of an underweight baby, whatever the gestation at delivery. These underweight babies are at risk of several early life complications such as hypothermia, low blood- sugar, feeding difficulties and viral infections. Aim of proposed system is analyse and predict risk during pregnancy physiological parameters like blood pressure, blood glucose level and weight whose changes during pregnancy can lead to complications can be identified and further complications induced from these changes can be prevented[6].

II. LITERATURE REVIEW

Pregnancy is a delicate stage requiring utmost care and medical attention towards health, as several abnormalities are induced in pregnant women health throughout this period. Health monitoring during pregnancy period can control, manage maintain and prevent abnormalities induced in physiological parameters, which may otherwise lead to severe complications leading to maternal and fetal mortality if unnoticed. The present research aims at catering health monitoring to women during pregnancy period to protect them from possible future complication in health.

(A) Aparna Gorthi, focused on the [1] goal of routine antenatal consultations is to predict and detect early complications of pregnancy allowing better management and hence outcome for both the mother and fetus. Antenatal check-ups currently comprise:

monthly clinical visits from third month of pregnancy, laboratory tests and obstetric ultrasounds at one each trimester. High risk pregnancies include women with history of complicated pregnancies and deliveries, diseases such as diabetes, hypertension, immunologic disorders, and pregnancies presenting with anomalies like malnutrition, obesity, intrauterine growth restriction, etc. Such cases require more frequent monitoring and specialist care. Use the methods in machine learning classification algorithm is decision tree based learning most closely captures the domain experts process of evaluation of a clinical case and as result A training accuracy of 72.4% was achieved with the 200 cases. The classifier reported results on the test cases with an accuracy of 65.5%.

(B) Lakshmi.B.N, said that [4] Pregnancy is a delicate stage in women life monitoring during pregnancy period can control, manage maintain and prevent abnormalities induced in physiological parameters.C4.5 provides great, precise, reliable and accurate prediction and an accuracy percentage of 98%. Hence, hybrid approach involving data transformation, C4.5 classifier application, rule extraction and range replacement gives better performance, accurate risk prediction, valid set of rules for further implementation and reliable solution to the complications faced by pregnant women during pregnancy.

III. PROBLEM DEFINITION

In India and in the other countries larger number of the people lives in the rural areas. They do not have sufficient knowledge as well as awareness regarding maternal health. As a result the maternal mortality rate is increasing day by day in our country. There are many reasons that causes behind this such as malnutrition, infections during pregnancy, anaemia, unawareness and lack of proper knowledge etc. It has been seen that in the country Bangladesh which has the highest rate of malnutrition which contributes in maternal mortality and low birth weight highly [7]. Around 30-40% of Bangladeshi babies weighing less than 2,500 g at birth which is among the highest rates of low birth weight in the world [8].It has been thought that improved nutrition for mothers during pregnancy would leave an impact to reduce this rate [7], [8]. If the mother can be guided with proper dietary advice from expert dieticians, it will help them and the fetus to stay healthy. Another most important thing which can prevent the negative outcomes of pregnancy is prenatal care. It has been observed that ,most of the women do not get proper recommendations during their pregnancy. It is not possible for all expecting mothers to visit clinics for every single complication they are passing through. In the rural areas, the expecting mothers need to relay on social taboos and other women of their neighbourhood. Their economic bindings, unwillingness and unawareness of visiting the doctors regularly maintaining proper time interval result in huge complications in their delivery. If they can be provided with proper medical guideline and reminder about their health conditions, the possibility of maternal health risk can be reduced. A continuous monitoring of the health condition of the expecting mother is necessary for this purpose. In most of the cases people do not have enough knowledge about the milestones of the development of the fetus (Figure 1). They do not know the symptoms which they are supposed to face at that

particular pregnancy stage. If any system can be developed letting them know about these things simply staying at home, they will also become conscious about their health. Apart from this, visiting the clinical experts for minor reasons is a hassle for the pregnant mothers.

Considering all those situations in developing countries like ours, an intelligent system is needed which will help the pregnant mothers during different stages of their pregnancy. This has motivated us to design a framework for a cloud based smart assistant for the expecting mothers. It will help our mothers to understand the information very easily. The objectives of providing such assistant to the expecting mothers are

- To assist the rural women who are not highly educated
- To reduce the number of prenatal visits for minor issues to the doctors
- Prepare the parents of the baby by estimating the delivery date
- Make them knowledgeable of all the stages of pregnancy and also making aware of the complications that may occur.
- Providing a perfect low cost personalized dietary guideline to reduce the malnutrition
- A continuous monitoring assessment of the mother's health to reduce complications

There is increased possibility of restricted fetal growth and birth of an underweight baby, whatever the gestation at delivery. These underweight babies are at risk of several early life complications such as hypothermia, low blood-sugar feeding difficulties and viral infections.

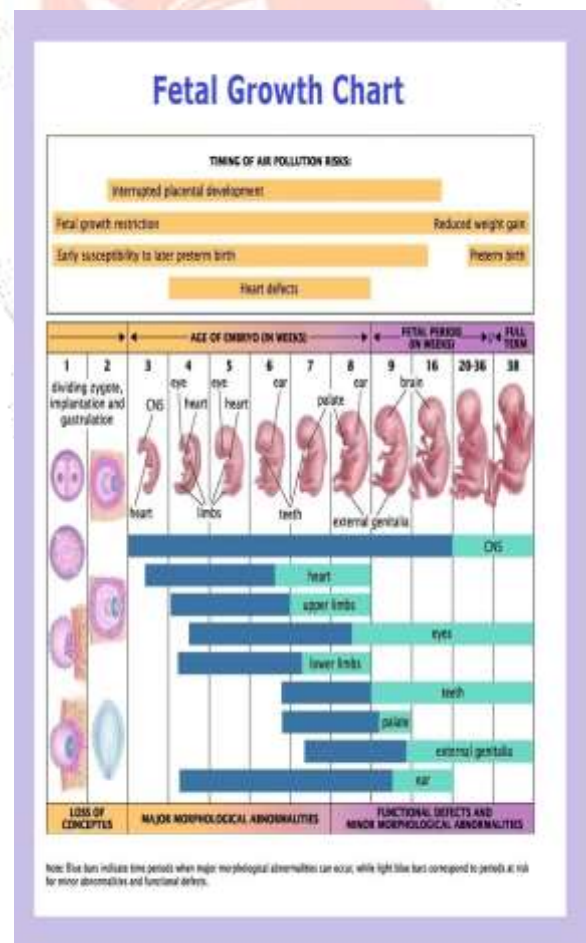


Figure 1 Growth Development of a Fetus

IV. PROPOSED WORK

Aim of proposed system is to analyse and predict risk during pregnancy physiological parameters like blood pressure, blood glucose level and weight whose changes during pregnancy can lead to complications which can be identified and further complications induced from these changes can be prevented. Considering all those situations in developing countries like ours, an intelligent system is designed which will help the pregnant mothers during different stages of their pregnancy. This is a framework for a cloud based smart assistant for the expecting mothers. It will help our mothers to understand the information very easily. The proposed hybrid approach bestows risk predictions based on values of a set of physiological parameters obtained from pregnant women thus providing them with an opportunity to control, manage, maintain and prevent probable severe health complications thus catering pregnant women with self-health care services by the help of the machine learning algorithm by storing the days of the mother in the database.

A. Dataset Collection: A Dataset collect from pregnant women like blood pressure, heart rate, gestational diabetes, preeclampsia .A dataset to train the classifier need to be prepared. which is being measured by using the adruino where the various sensors for measuring the blood pressure, heart rate, mother weight which increase with the month of the pregnancy.(Fig2, Fig3, Fig4)

B. Preprocessing: The first step of pre-processing is all about correcting the problems in data before building a machine learning model using data. Problems can be of many types like missing values, attributes with a different range.

C. Testing By representing report for testing it gives correct result or prediction.

E. Risk Prediction Machine learning algorithm is used to predict result. It applies on the dataset and gives prediction in the form of risk (normal, high, very high). Machine learning is the subfield of computer science that gives computers the ability to learn without being explicitly programmed. Evolved from the study of pattern recognition and computational learning theory in artificial intelligence, machine learning explores the study and construction of algorithms that can learn from and make predictions on data. Machine learning is closely related to computational statistics, which also focuses in prediction-making through the use of computers.

The mother's required data is collected and is stored in the database and when the mother is being done with the routine check-ups first the data in the database is being checked and according to which it been observed about the changes that is occurring in the mother and also make aware of the complications if any if there some abnormalities occur.

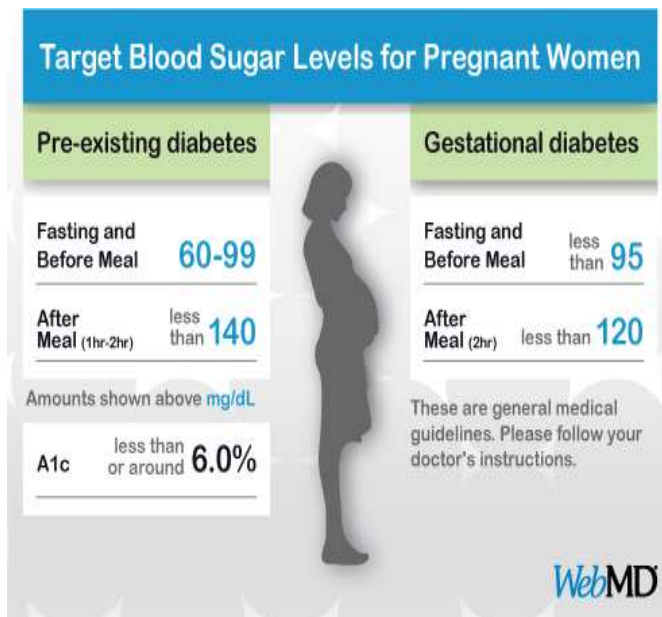


Figure2 Sugar Levels of a Pregnant Lady

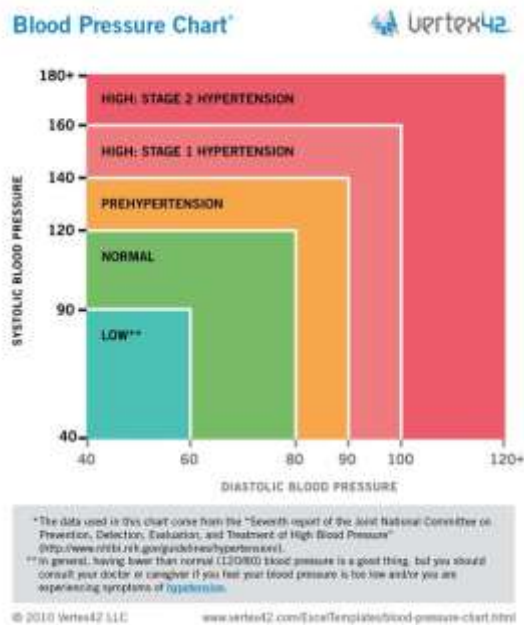


Figure3 Blood Pressure Chart

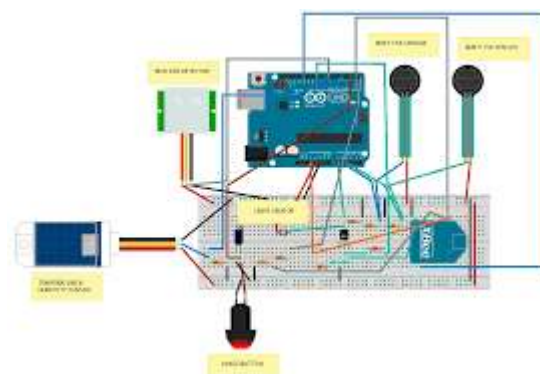


Figure4 Blood Pressure and Sugar level Measure

V. ALGORITHM IMPLEMENTED-C4.5 ALGORITHM

C4.5 builds decision trees from a set of training data in the same way as ID3, using the concept of information entropy. The training data is a set of already classified samples. Each sample consists of a p-dimensional vector, where the represent attribute values or features of the sample, as well as the class in which falls.

At each node of the tree, C4.5 chooses the attribute of the data that most effectively splits its set of samples into subsets enriched in one class or the other. The splitting criterion is the normalized information gain (difference in entropy). The attribute with the highest normalized information gain is chosen to make the decision. The C4.5 algorithm then recurs on the smaller sub lists.

This algorithm has a few base cases.

- All the samples in the list belong to the same class. When this happens, it simply creates a leaf node for the decision tree saying to choose that class.
- None of the features provide any information gain. In this case, C4.5 creates a decision node higher up the tree using the expected value of the class.
- Instance of previously-unseen class encountered. Again, C4.5 creates a decision node higher up the tree using the expected value.

ALGORITHM

The general algorithm for building decision trees is:^[11]

1. Check for the above base cases.
2. For each attribute a , find the normalized information gain ratio from splitting on a .
3. Let a_{best} be the attribute with the highest normalized information gain.
4. Create a decision *node* that splits on a_{best} .
5. Recur on the sublists obtained by splitting on a_{best} , and add those nodes as children of nodes.

The algorithm used is the one pass SVM algorithm which is the most popular algorithm and gives more appropriate information than C4.5 algorithm. C4.5 are very useful for machine learning. A support vector machine is a classification method used in datasets classification and regression. This is a non-linear classification algorithm and will help to perform data mining, text mining and pattern recognition. It is often said as better classification results over to other classification algorithms like C4.5, Decision Tree etc. It delivers good and precise solution to optimal dataset problems.

Support Vector Machines (SVMs) are one of the latest development of supervised machine learning technique.[12]. Though SVM were known since late seventies it started to receive attention on late nineties. It was applied basically to pattern recognition, also used for

pattern classification problems like image recognition, text recognition, etc. However many research was implemented based SVM in solving intrusion detection problem such as in SVMs works mainly by deriving a hyper plane that maximizes the separating margin between two classes. The feature vectors that lie on the boundary of separation vectors are called support vectors. SVM are fantastic because they are very resilient to over fitting.

VI CONCLUSION

Pregnant women undergo several problems like blood pressure, heart rate, temperature increase and decrease, gestational diabetes and hypertension during stage of pregnancy. It leads to complication in pregnancy which may result to mortality of mother and fetus so we need to protect pregnant women. From this survey, C4.5 classification algorithm has greatest potential in accuracy for predicting the risk levels during pregnancy. With its privileges and features, takes the responsibility to provide the initial hand of care of pregnant women. We believe only a healthy mother is able to give birth to a healthy infant. The application, even though created to focus the user base at rural areas, also posses the ability to help out any women who is expecting a baby. The application has been tried to be kept as simple as possible and the user interface will be developed based on expert's advice. The simplicity in making will be ensured so that the application can remain a generalized solution for all our women. Integration of more languages will be able to support a greater community as well. A noble cause for a noble community; this proposed system will also inspire people to come up and implement better services and ensure the safety of the entire human race as a whole.

ACKNOWLEDGMENT

We wish to acknowledge our guide Er. Jinu Thomas for all the support and guidance she has administered for this research and also helped us with all other contributors that helped us to make this idea a viable application.

VII REFERENCES

- [1] V.Srinivasan,G.Rajenderan, J.Vandar Kuzhali, M.Aruna,"Fuzzy Classification to Identify the Risk in Diabetic Pregnancy"Conference IEEE 2011.
- [2] Lakshmi.B.N , Dr.Indumathi.T.S , Dr.Nandini Ravi " Prediction Based Health Monitoring in Pregnant Women" International Conferences on Emerging in Engineering, Science and Technology(ICETEST-2015).
- [3] Jostinah Lam ,Yudha Aditya Noor ,Eko Supriyanto, " Ontology Driven Knowledge Base for High Risk Pregnancy Management" 2015 4th International Conference on Instrumentation, Communications, Information Technology, and Biomedical Engineering (ICICI-BME) Bandung, November 2-3,2015
- [4] Lakshmi.B.N ,Dr.Indumathi.T.S ,Dr.Nandini Ravi," An Hybrid Approach for Prediction Based Health Monitoring in Pregnant Women" International Conference on Emerging Trends in Engineering, Science and Technology (ICETEST - 2015).
- [5] Lakshmi.B.N, Indumathi.T.S,Nandini Ravi," A Comparative Study of Classification Algorithms for Predicting Gestational

Risks in Pregnant Women”International Conference on Computers, Communications, and Systems 2015.

- [6] Lakshmi.B.N, Indumathi.T.S,Nandini Ravi, “A study on C.5 Decision Tree Classification Algorithm for Risk Predictions during Pregnancy” International Conference on Emerging Trends in Engineering, Science and Technology (ICETEST) 2015.
- [7] A. Faruque, A. S. Ahmed, T. Ahmed, M. M. Islam, M. I. Hossain, S. Roy, N. Alam, I. Kabir, and D. A. Sack, “ Nutrition: basis for healthy children and mothers in bangladesh,” Journal of Health, Population and Nutrition, pp. 325– 339, 2008.
- [8] M. Basher, S. Kabir, S. Ahmed, M. Miah, and M. Kamal, “ Prenatal nutrition among rural bangladeshi pregnant women.,” Mymensingh medical journal: MMJ, vol. 20, no. 4, pp. 548– 557, 2011.
- [9] Google Search, www.google.com, complication during pregnancy
- [10] Google Search, www.google.com, c4.5 Algorithm.
- [11] S.B. Kotsiantis, Supervised Machine Learning: A Review of Classification Techniques, Informatics 31(2007) 249-268, 2007.
- [12] Andrew Ng. Cs229 lecture notes, Autumn 2014

