31st Annual International Conference of the IEEE EMBS
Minneapolis, Minnesota, USA, September 2-6, 2009

Automated risk assessment tool for pregnancy care

Aparna Gorthi†, Celine Firtion† and Jithendra Vepa, *Member, IEEE*

***Abstract*—Clinical decision support systems augment the quality of medical care by aiding healthcare workers in the evaluation and management of complicated cases. Clinical decision support systems are especially instrumental in quickly assessing the criticality of pregnancy as it involves interpreting multiple maternal and fetal parameters. We propose a machine learning approach for early determination of the risk category of pregnancy based on patterns gleaned from profiles of known clinical parameters. In particular, we demonstrate the usefulness of classification and regression trees in solving multivariate problems in obstetric care since the decision making process and the importance of specific parameters are clearly illustrated in the tree. As proof of concept, an application use case has been presented.**

***Keywords*—** pregnancy risk assessment, decision tree-based learning, clinical decision support system (CDSS).

I. Introduction

PREGNANCY is a period of risk for both the mother and the fetus. In developing countries like India, it results in very high maternal (300/100000 live births compare to 0.053/100000 live births in the UK) and neonatal mortality rates (44/1000 live births compared to 0.28/1000 live births in the UK) [1]. One of the main reasons for these high rates in developing countries is the lack of trained specialists (6 doctors/10000 in India as compared to 23 doctors/10000 in UK) [2]. To decrease these rates, early diagnosis of high-risk pregnancies and quick referrals are of paramount importance.

The main 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 (one each trimester) [3]. High risk pregnancies include women with history of complicated pregnancies and/or deliveries, diseases such as diabetes, hypertension, immunologic disorders, and pregnancies presenting with anomalies like malnutrition, obesity, intra-uterine growth restriction, etc. Such cases require more frequent monitoring (Doppler ultrasound, fetal monitoring, etc.) and specialist care. The interpretation and categorization of data from clinical and diagnostic investigations is done by obstetricians/physicians or healthcare workers based on their knowledge and know-how. Currently the process is highly subjective and requires interpretative skills.

Clinical decision support systems (CDSS) have emerged as an important component of technology-assisted medical care by helping various clinical processes through the use of experiential knowledge. CDSS can also enhance the decision making process of healthcare professionals by incorporating clinical guidelines in their rule engine. Such systems also allow the extraction of meaningful information from high-dimensional data sets with a priori feature elimination. CDSSs have been developed since the 1970s, but very few are related to pregnancy. Existing automated tools are available only for fetal heart rate analysis, screening for Down syndrome, gestational diabetes and labour management [4-7].

In this paper, we present a model based on decision tree leaning that captures the decision making process of a medical professional in determining the risk associated with a pregnancy. A wide range of parameters have been identified to play a critical role in evaluating the clinical case. A parameter selection approach is presented to assist physicians in identifying the clinical parameters best suited for a particular discrimination task.

Of all the methods in machine learning, decision tree-based learning most closely captures the domain expert's process of evaluation of a clinical case. A decision tree is a predictive model to represent various decisions or rules and their possible outcomes (class labels). Decision tree algorithms like ID3 [8] have been successfully applied to various medical diagnostic problems in oncology, liver pathology, diagnosis of thyroid diseases etc [9]. We present a solution that uses classification and regression trees to evaluate the risk category of a pregnancy. Section II describes the overall approach and methodology adopted in building such a model. Sections III and IV reflect on the results from the experiments conducted and propose an application use case for the automated pregnancy risk assessment model.

II. Methods

A. System workflow

Our goal was to develop a CDSS which provides the healthcare worker with a simple pregnancy risk scale and guide him/her in the decision-making process during evaluation of the case. The approach followed included (Fig. 1): 1) Use synthetically generated clinical cases classified as normal, moderate or high risk pregnancies by an obstetrician to extract relevant parameters for risk assessment, 2) Employ a suitable machine learning algorithm to extract risk trends from the training data and predict risk potential of a given pregnancy.

†Both authors contributed equally to the work.

Aparna Gorthi is with Philips Research Asia-Bangalore, Bangalore, India 560045 (phone: +91 80 4189 2264; fax: +91 80 4189 2265; e-mail: aparna.gorthi@philips.com).

Celine Firtion is with Philips Research Asia-Bangalore, Bangalore, India 560045 (email: celine.firtion@philips.com)

Jithendra Vepa is with Philips Research Asia-Bangalore, Bangalore, India 560045 (email: vepa.jithendra@philips.com)

978-1-4244-3296-7/09/$25.00 ©2009 IEEE

B. Selection of medical parameters

Literature survey was carried out to extract the most relevant maternal and fetal parameters that play a significant role in affecting maternal and fetal health and those used in routine care [2] [10]. The range of values observed under each category for the parameters selected were also defined based on literature and clinical guidelines. A total of 31 features were finally selected for the analysis. These parameters are routinely recorded during the follow-up of normal and high-risk pregnancies and thus would be most suitable for training our classifier. These parameters are listed in Table 1.

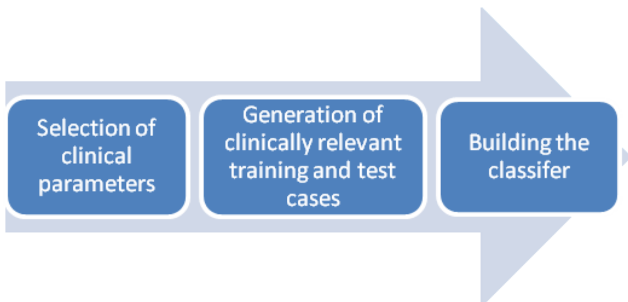

Figure 1. System Workflow

C. Generation of clinical cases

Due to the lack of availability of real patient records, clinically meaningful cases under each risk category were generated using a rule based engine:

1. Based on the information collected from the clinical guidelines, the lower and upper bounds of acceptable values were defined.
2. Values for independent parameters like albuminuria, metrorragia, etc. were generated using a Gaussian distribution of the acceptable value range assuming a standard deviation that was clinically relevant.
3. For parameters which were inter-dependent (like systolic and diastolic blood pressure, gestational age and weight gain), a Gaussian for the independent parameter was first produced with a meaningful standard deviation (Fig 2). Values of the dependent variables were then randomly generated within the constraints defined by the selected value of the independent variable.
4. Each such case comprising all the parameters was manually validated as clinically meaningful and representative of the designated risk category by a trained obstetrician.

A total of 200 training cases and 40 test cases were generated.

D. Classification Algorithm

We have used classification and regression trees (CART) [10] since they can simulate the clinical decision rules. CART can deal with incomplete data and different types of data (both numerical and categorical data). Classification tree has been used to determine the class to which the clinical profile of a pregnant woman belongs to. The generated decision tree was also used to identify the critical parameters in the decision making process.

Table 1 Parameters used in Classification

(Features listed below are divided into two subcategories for each class – parameters in shaded gray were found to be significant; *Parameters in italics were removed after feature elimination experiments discussed in Section III)*

CLASS	PARAMETERS USED IN AIDING RISK ASSESSMENT
Routine Clinical parameters	Maternal Age, Gravidity, Gestational Age, Height, Weight, Weight gain, Blood Pressure, Metrorragia, Uterine Contractions, Amniotic Fluid Loss. Seizures
	Parity, Ringing in the ears, Headache, Change of vision, Epigastric pain, Edema
Laboratory parameters	Albuminuria, hyperglycemia
Fetal monitoring related parameters	Basal fetal heart rate, Variability, Accelerations, Decelerations
Ultrasound and Doppler parameters	Fetal Movement, Placental Localization, Growth of Fetus, Umbilical Doppler
	Fetus presentation, uterine and fetal Doppler signals (cerebral, umbilical), amniotic fluid index)

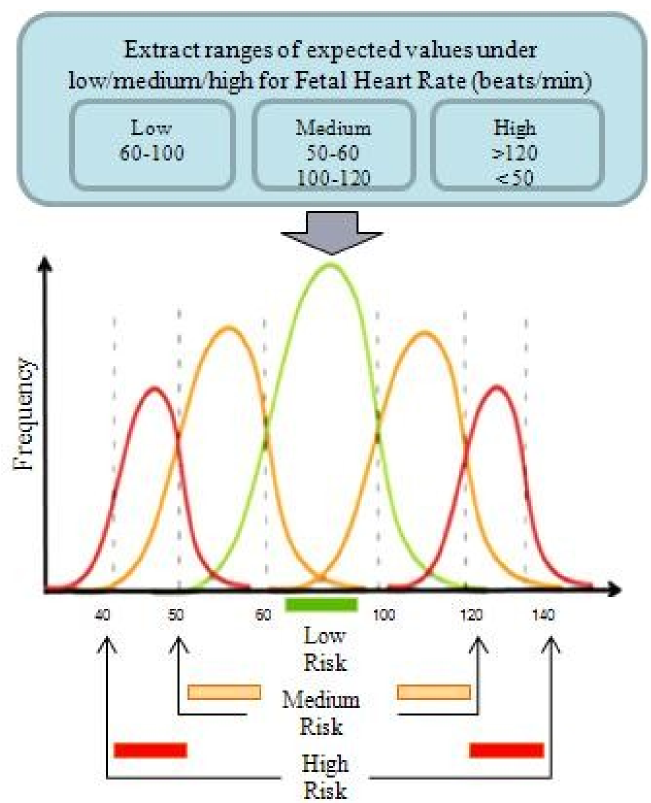

Figure 2. Selection of parameter values under each risk category

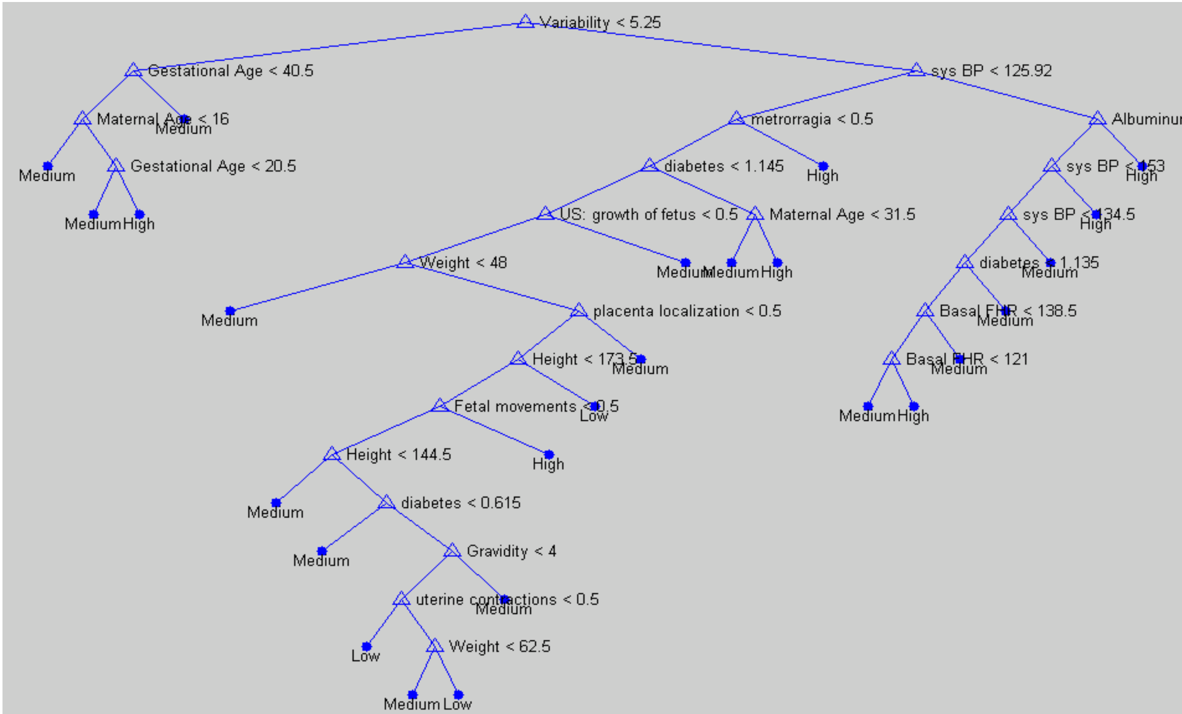

Figure 3. Decision Tree

III. RESULTS

Two hundred cases and their known class labels were used to train a classification tree. The created decision tree is shown in Fig. 3. Trials of training the classifier with different subsets of the training samples yielded similar decision trees indicating significance of the learning. In order to evaluate the importance of each parameter, recursive feature elimination experiments were conducted. Out of the 31 features used initially only 22 were found to be involved in the classification of a clinical case and hence the other 9 features were eliminated in further experiments (Table 1).

Table 2 Confusion Matrix

Risk	Low	Medium	High
Low	60	2	2
Medium	4	64	2
High	2	3	63
Total No. of Cases	66	69	67
% Mis-classification	9.1	7.2	5.9

In order to evaluate the classification efficiency, five-fold cross validation experiments were carried out. A training accuracy of 93.4% was achieved with the 200 cases. The classifier reported results on the test cases with an accuracy of 82.5%. (Table 2).

An analysis of how the classifier performed in evaluating test cases under each of the three categories is illustrated in the confusion matrix (Table 2). For each of the three classes, a unique set of parameter values clearly defined the boundary conditions. As shown in the confusion matrix, the rate of mis-classification is low.

IV. DISCUSSION

Computer-based diagnostic pregnancy care systems have been developed [11] in the past; however most of them are confined to labour management [12-16]. The most popular and routinely used system for the management of high-risk pregnancies is FetalCare™ from the Sonicaid™ which identifies abnormal fetal heart rate traces during pregnancy and shortens the monitoring time by half. It is based on electronic archive of health records and analysis of results initially with 8 000 cases for the first algorithm and the latest version uses 73 802 cases [12]. Others have developed automatic interpretation of the fetal heart trace based on medical guidelines and have associated it with a telemonitoring system [13] or to other investigations like ultrasound based biophysical profile [6].

The CDSS presented in this paper focuses on antenatal care by addressing maternal and fetal health during pregnancy. The assessment of severity of a case is strengthened by providing a simple scaled interpretation (low, medium and high risk) which can be understood even by non-specialists. If necessary, the system can give detailed explanation on the parameters used for classification.

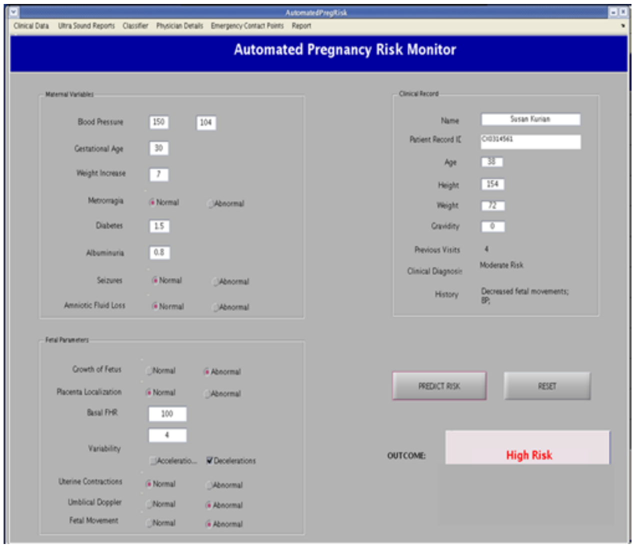

Figure 4 Screenshot of Automated Pregnancy Risk Assessment Tool

The classifier was constructed through decision tree-based learning performs reasonably well given the limited sample size. The classifiers were trained using different training data sizes and the leaned decision trees exhibited consistency in all cases. Decision tree learning is particularly suitable in the domain of medical informatics since clinical experts are familiar with the symbolic representation of the decision making process in the tree form and hence they can assess the outcome predicted by the classifier better.

Overall, the order of importance of the parameters involved in decision making extracted from our model fall in accordance with the clinical relevance of these parameters thus validating the methodology. Fetal heart rate variability measurement, though not a routine investigation, tends to take precedence over other parameters and dictate the management when an abnormal pattern of the trace is found. The other parameters which are prominent in the decision tree are also those which take precedence for the clinician's decision; indeed these parameters, when abnormal, put the mother and/or fetus at high-risk of complications: high maternal blood pressure and presence of albuminuria (hypertension and preeclampsia diseases), gestational age (post-term pregnancy), maternal age (teenage pregnancy), metrorragia (bleeding due to placenta praevia, abruption placenta, etc.), hyperglycemia (diabetes), abnormal fetal growth (intra-uterine growth restriction, chromosomal anomalies), etc.

The model can be augmented to perform at clinically acceptable accuracy levels (>98%) with a larger number of test cases and additional rules. Fig. 4 shows the screenshot of an easy to use graphical interface that allows the physician to enter data as it is being collected and assess the risk of pregnancy based on the current set of values and previous history. The tool can be improvised to also store all the clinical details of the patient's previous visits and use this information in the analysis.

V. CONCLUSION

We propose the use of the decision tree-based classification system as an automated risk assessment tool where the parameters are entered into the system either manually by the physician or healthcare worker through a simple software interface (clinical parameters, and some of the paraclinical investigations like urine test result) or automatically extracted from devices like ultrasound machines and fetal monitoring devices. As all the antenatal parameters will be systematically checked and analyzed, the decision tree-based system would offer the following benefits:

1. Comprehensive and systematic follow-up allowing early detection and referral of high-risk complications during pregnancy. This would be especially useful for patients living in remote areas where the healthcare worker might not be a specialist.
2. Decreased risk of errors and medico-legal liability
3. The tool could be integrated with patients' database and thus increase the number of patients accessible with tele-transmission of data, allowing effective planning and utilization of resources.

REFERENCES

[1] The state of the world children 2009-Maternal and newborn health. *UNICEF*, 2008

[2] World Health Statistics 2008. *WHO Press*, Geneva, 2008.

[3] National Collaborating Centre for Women's and Children's Health, "Antenatal care, routine care for the healthy pregnant woman" in *Clinical Guidelines*. *RCOG Press*, 2003.

[4] A A-Bentanzos, B G- Berdiñas, V M-Bonillo and S L- Gonźalez, "The NST-EXPERT project: the need to evolve." in *Artif Intell Med.*, vol. 7(4), pp. 297-313,1995.

[5] N.J. Wald, H.C. Watt, and A.K. Hackshaw, "Integrated screening for Down's syndrome based on tests performed during the first and second trimesters." in *New England Journal of Medicine*, vol. 341(7), pp. 461-467, 1999.

[6] M.E. Hernando, E.J. Gómez, F. d Pozo and R. Corcoy, "DIABNET: a qualitative model-based advisory system for therapy planning in gestational diabetes." in *Med Inform*, vol. 21(4), pp. 359-74, 1996.

[7] R.D. Keith and K.R. Greene," Development, evaluation and validation of an intelligent system for the management of labour." in *Baillieres Clin Obstet Gynaecol,* vol. 8(3), pp. 583-605, 1994.

[8] A Colin, "Building decision trees with the ID3 algorithm", *Dr. Dobbs Journal,* 1996.

[9] I Kononenko, "Machine learning for medical diagnosis: history, state of the art and perspective", *Artificial Intelligence in Medicine*, vol. 23(1), pp. 89-109, 2001.

[10] L. Breiman., "Classification and regression trees", *Chapman & Hall*, Boca Raton, 1993.

[11] WHO-Department of reproductive health and research, "Integrated management of pregnancy and childbirth-managing complications in pregnancy and childbirth: A guide for midwifes and doctors." in *WHO Press*, Geneva, 2000.

[12] Huntleigh Healthcare Ltd., "Sonicaid fetalcare antepartum analysis. Clinical application guide." Issue 1.3, Cardiff 2005.

[13] Philips Electronics North America Corporation, "OB TraceVue-Automatic CTG analysis and alerting-Application note.", 2003.

[14] A. A-Betanzos, V. M-Bonillo and C. H-Sande, "Foetos: an expert system for fetal assessment." in *IEEE Trans Biomed Eng.* Vol. 38(2), pp. 199-211, 1991.

[15] R. D. Keith, S. Beckley, J.M. Garibaldi, J.A. Westgate, E.C Ifeachor and K.R. Greene. "A multicentre comparative study of 17 experts and an intelligent computer system for managing labour using the cardiotocogram." in *Br J Obstet Gynaecol,* vol 103(5), pp. 489-90, 1996.