International Journal of Scientific Research in Computer Science, Engineering and Information Technology
© 2017 IJSRCSEIT | Volume 2 | Issue 4 | ISSN : 2456-3307

Naive Bayesian Model for Predicting Mode of Delivery

Dr. K. Lakshmi*1, Dr. P. Raja Priya2, Dr. T. Nivedhitha3

*1Computer Science and Engineering, Periyar Maniammai University, Thanjavur,Tamilnadu, India
2,3Department of Obstetrics and Gynecology, Pondicherry Institute of Medical Sciences, Pondicherry, India

ABSTRACT

Giving birth to a child gives anxiety to any family due to many uncertainties even at the technology era. The research work analyzes the data of pregnant women and brings out a model for predicting the mode of delivery of baby. This paper presents the analytical data & effectiveness of the Bayesian model. We have considered the features such as age, BMI (Body Mass Index), GDM (Gestational Diabetes Mellitus) etc., The purpose of this study was to determine the relative effects of pregnancy and mode of delivery.

Keywords: Data Classification, Mode of Delivery, Prediction, Pregnancy, Bayesian Model.

I. INTRODUCTION

Since 1900, trends have changed dramatically with the increase of technology and modern medical practices. Medical technology has made childbirth a much safer experience over the past century for both mother and baby.

In clinical obstetrics, vaginal deliveries are usually preferred to C-sections in low-risk pregnant and laboring women, because C-sections have been associated with an increased risk of complications to anaesthesia, excessive blood loss, respiratory complications, longer recovery periods in the mother, risks associated with C-section antea in subsequent pregnancies [1].

There may be different kinds of Childbirth and delivery methods:

Normal Vaginal Delivery (NVD) - In a vaginal birth, the baby is born through the birth canal. Most women give birth at around 38-41 weeks of pregnancy.

Cesarean Section (LSCS) is the delivery of a baby through a surgical incision in the mother's abdomen and uterus.

Vacuum Extraction - A vacuum extraction is a procedure sometimes done during the course of vaginal childbirth. During vacuum extraction, a health care provider applies the vacuum (a soft or rigid cup with a handle and a vacuum pump) to the baby's head to help guide the baby out of the birth canal.

Forceps Delivery - A forceps delivery is a type of operative vaginal delivery. It's sometimes needed in the course of vaginal childbirth. In a forceps delivery, a health care provider applies forceps (an instrument shaped like a pair of large spoons or salad tongs) to the baby's head to help guide the baby out of the birth canal.

Pre Term Vaginal Delivery (PTVD) – When a delivery occurs after 28 weeks and before 37 weeks Period of gestation.

Although all the modes of delivery are common, vaginal delivery without any instrumentation and after 37 weeks is considered ideal. When there are few complications involved other mode of delivery has to be opted. There are certain risks and complications associated with a caesarean delivery. Additionally, it involves prior preparation such as funding and pain control. Thus, based on logistics and medical parameters, choosing the best method is important in successful childbirth of the pregnancy. By monitoring the entire period of nine months of gestation, many data are collected. Use of appropriate cleaning, sorting and classification techniques reveals patterns which can predict possible threats or anomalies. The objective of our study is to predict the mode of delivery based on specific parameters identified separately from amongst 36 parameters available in various test reports. Such an

early prediction would help women to be mentally and financially prepared.

We propose to use data mining technique Naïve Bayesian method for finding the probability of mode of delivery given the medical and physical conditions.

This paper discusses the data set and the features used in the model preparation in section II. In section III, Bayesian model is explained; also this section covers the training phase and prediction phase details. Finally section IV discusses the data distribution and the accuracy level of Bayesian model on the data taken.

II. DATA SET

We have collected data of pregnant women from a hospital. The data collected consisted of the following parameters, Patient ID no, name, age, gravida (Number of times the woman has become pregnant including the present pregnancy), pre mature birth, living children, abortion, Period of gestation (POG) in week, height in meter, weight in kg, body mass index (BMI), blood pressure category of body based on BMI, POG at abortion, threatened preterm (when there is a situation in which there are few signs which shows that the baby may be delivered before 37 weeks), missed abortion, incomplete abortion (process of abortion has started but not completed), complete abortion (the process of abortion has started and has completed), Gestational Diabetes Mellitus (GDM- when the woman is diagnosed to be diabetic first time after conceiving), PIH (PIH- when the BP is >140/90 diagnosed for the first time in pregnancy after 20 weeks POG), preeclampsia (when BP is >140/90 and liver enzymes are deranged), eclampsia (when patient has seizure due to increased BP in a woman who is not already a known case of seizure disorder, sometimes can occur in normal BP) age_at_delivery_weeks, Preterm Premature Rupture of the Membranes (PPROM - when the amniotic sac ruptures before 37 weeks POG), mode_of_delivery, post partum haemorrhage (PPH - when patient bleeds excessively after the delivery of the baby and placenta), APGAR-1_min, 5_min (Parameters used to assess the condition of the baby immediately after delivery), birth_weight, viability_at_birth, sex, special_risk, indication (in case of instrumentation and LSCS the reason for which it is done).

We have collected more than 220 patient records. For our experimentation we have considered only important attributes like Age, Category of body type based on the Body Mass Index, Gestational Diabetes Mellitus at Period of Gestation at Diagnosis (GDM), Preterm Premature Rupture of the Membranes (PPROM), Pregnancy Induced Hypertension (PIH) as discussed in [4] [5]. BMI is calculated based on the woman's height and weight, so we have not considered woman's height and weight separately. Moreover there are many features with most of the values empty. For example abortion_threatened_weeks is applicable only for abortion cases. So for most of the patient this data is irrelevant, since abortion cases are very less in the data set collected. Hence we have considered only five features are considered for predicting the delivery mode.

Wherever data is missing, data are filled with most probable values for that feature. We have not considered cases of Abortion here. Wherever data is missing for many of the attributes those records are not considered. We use the patient details who have given successful child birth for building the Bayesian Model.

III. BAYESIAN MODEL

Implementation of Mining Model includes various algorithms and techniques such as Classification, Clustering, Regression, Artificial Intelligence, Neural Networks, Association Rules, Decision Trees, Genetic Algorithm, Nearest Neighbor method *etc*. Classification is one of the most frequently studied problems by data mining and machine learning (ML) researchers.

It consists of predicting the value of a (categorical) attribute (the class) based on the values of other attributes (the predicting attributes). In the present study we use the Bayesian Classification algorithm. Bayes classification has been proposed that is based on Bayes rule of conditional probability. Bayes rule is a technique to estimate the likelihood of a property given the set of data as evidence or input.

The approach is called "naïve" because it assumes the independence between the various attribute values. Naïve Bayes classification can be viewed as both a descriptive and a predictive type of algorithm. The probabilities are descriptive and are then used to predict the class membership for a target tuple. The naïve Bayes approach has several advantages: it is easy to use; unlike other classification approaches only one scan of the training data is required; easily handle mining value by simply omitting that probability [3]. An advantage of the naive Bayes classifier is that it requires a small amount of training data to estimate the parameters (means and variances of the variables) necessary for classification. Because independent

variables are assumed, only the variances of the variables for each class need to be determined and not the entire covariance matrix. In spite of their naive design and apparently over-simplified assumptions, naive Bayes classifiers have worked quite well in many complex real-world situations.

The Naïve Bayes classifier [2] selects the most likely classification V_{nb} given the attribute values $a_1, a_2, a_3 \ldots a_n$.

$$V_{nb} = argmax_{vj \in V} P(v_j) \Pi P(a_i|v_j) \quad (1)$$

$P(a_i|v_j)$ is estimated using

$$P(a_i|v_j) = \frac{n_c + mp}{n+m} \quad (2)$$

Where:

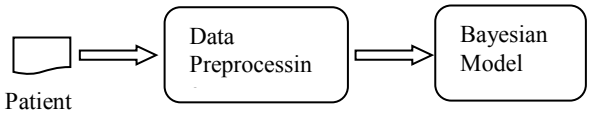
n = the number of training examples for which $v = v_j$

n_c = number of examples for which $v = v_j$ and $a = a_i$

p = a priori estimation for $P(a_i|v_j)$

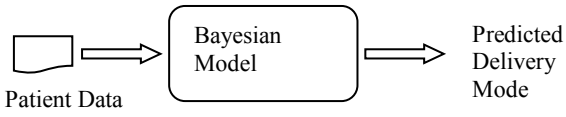
m = the equivalent sample size

A. Training Phase



During the training phase the data of the patients are fed into the system. Unwanted records & features are deleted. Probability values of each delivery mode are calculated and conditional probability of attributes and delivery modes are calculated as per the equation (2).

B. Prediction of Delivery



Once the probabilities are calculated the new patient data can be given to Bayesian model to get the predicted delivery mode using equation (1).

IV. EXPERIMENT DATA ANALYSIS

We have 188 cases represented here. Mode of delivery legends used in this section as follows:

- EM_LSCS – Emergency LSCS
- EL_LSCS – Elective LSCS
- NVD – Normal Vaginal Delivery
- PTVD – Pre Term Virginal Delivery
- VACCUM – Vacuum delivery
- FORCEPS – Forceps delivery

First we give a general distribution of the data in Table I and Fig. 1. As we can see from Table I and Fig. 1, most of the deliveries are normal deliveries (NVD) only. On account of medical complications other types of delivery are conducted. Depending on the medical conditions the doctors take the decision of LSCS.

TABLE I. MODE OF DELIVERY AND THEIR COUNT

Mode of Delivery	Total
NVD	114
EM.LSCS	32
EL.LSCS	15
PTVD	13
VACUUM	12
FORCEPS	2

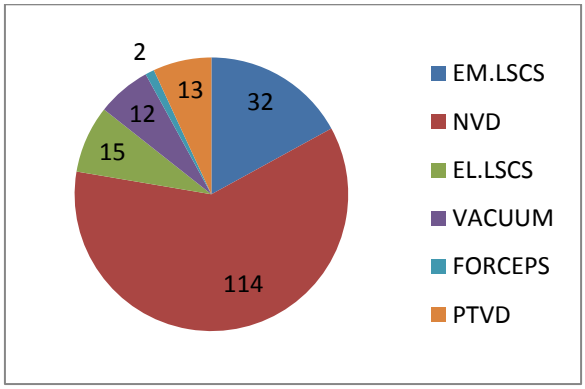


Figure 1. Mode of Delivery data distribution.

Table 2 gives the number of cases under various attributes with possible values and mode of deliveries. Attribute Age is numerical value which is categorized as Teen, Very Young, Young, Mid, Mid_Old and Old. Age less than twenty is labeled as Teen, other category labels are given with five years incrementally. Attribute Category is based on the Body Mass Index. As per the BMI body type is categorized as Normal, Obese, Over Weight and Under Weight. GDM and PIH take the value of Yes/No indicating whether the women developed diabetes or Hyper Tension during pregnancy. PPROM also takes the Yes/No values based on membrane rupture condition.

TABLE II

DISTRIBUTION OF DATA BASED ON THE VARIOUS ATTRIBUTES AND THEIR VALUES

Attributes	Attribute Values	Mode of Delivery					
		EM_LSCS	EL_LSCS	NVD	PTVD	VACUUM	FORCEPS
AGE	TEEN	4	0	11	0	1	0
	VERY YOUNG	8	4	55	7	5	0
	YOUNG	16	10	39	6	5	1
	MID	3	1	6	0	0	1
	MID_OLD	1	0	1	0	0	0
	OLD	0	0	2	0	1	0
CATEGORY	NORMAL	9	3	41	7	1	0
	OBESE	18	8	44	4	6	0
	OVER_WEIGHT	4	2	22	2	4	2
	UNDER_WEIGHT	1	2	7	0	1	0
GDM	YES	6	3	12	0	3	2
	NO	26	12	102	13	9	0
PIH	YES	5	0	11	1	2	0
	NO	27	15	103	12	10	2
PPROM	YES	2	1	0	7	1	0
	NO	30	14	114	6	11	2

We have built the Naïve Bayesian classifier as described in section III based on the values given in Table II. Prediction of delivery mode is done by using the posterior probability values calculated by the equation (1) given in section III. The training data itself is given as input to the Bayesian model. For example we give input as Age = Mid, Category = Normal, GDM= No, PIH = No, PPROM = No. Classifier gives the highest probable delivery mode as predicted delivery mode.

Out of 188 records, the naïve Baysian model is able to predict 118 records correctly. That means 62.76% accuracy.

Further we have analyzed the data to learn how much misclassification has happened with respect to specific delivery mode. All the EL_LSCS were misclassified as NVD. Except two, all other EM_LSCS are misclassified. Except three all other NVDs are classified as NVDs. Five PTVDs are classified correctly. None of the VACUUM and FORCEPs cases are identified correctly.

We can say the Naïve Bayesian model built is capable of predicting normal deliveries correctly. Other types of deliveries were not predicted correctly using the model we have built.

This model may be enhanced with additional knowledge or feature attributes. However the chosen attributes are not sufficient to give better prediction.

V. REFERENCES

[1]. Beth Ross (2017) Johnson Memorial Hospital [Online]. Available: http://blog.johnsonmemorial.org/blog/what-type-of-birth-is-right-for-you-and-your-baby

[2]. JiaWei Han, Micheline Kamber, Kian Pei, Datamining concepts and Techniques, Morgan Kaufmann Publishers 2012.

[3]. Alisha Kamat, Veenal Oswal, Manalee Datar, "Implementation of Classification Algorithms to Predict Mode of Delivery", International Journal of Computer Science and Information Technologies, Vol. 6 (5) , pp. 4531-4534, 2015

[4]. Nadia Moledina, Sheila McDonald, Donna Slater and Suzanne Tough, "Pre-pregnancy Body Mass Index (BMI) and delivery outcomes in a Canadian population", BMC Pregnancy and Childbirth 2014 14:422 DOI: 10.1186/s12884-014-0422-y

[5]. Michelle A. Kominiarek, Paul VanVeldhuisen, Judith Hibbard, Helain Landy, Shoshana Haberman, Lee Learman, Isabelle Wilkins, Jennifer Bailit, Ware Branch, Ronald Burkman, Victor Hugo Gonzalez Quintero, Kimberly Gregory, Christos Hatjis, Matthew Hoffman, Mildred Ramirez, Uma M. Reddy, James Troendle, Jun Zhang Am J Obstet Gynecol, "The Maternal Body Mass Index: A Strong Association with Delivery Route", American Journal of Obstetrics & Gynecology 2010 Sep; 203(3): 264.e1–264.e7. Published online 2010 Jul 31. doi: 10.1016/j.ajog.2010.06.024 PMCID: PMC2933947