**Predicting the mode of childbirth using Machine Learning**

**Abstract:**

Nowadays, the method of delivery is a crucial factor in ensuring the well-being of both mother and baby. Currently, the decision on the mode of delivery is usually made by the attending physician, but if the wrong method is chosen, it can lead to various short-term and long-term health problems for both the mother and baby. The number of cases where doctors unnecessarily suggest a cesarean delivery is on the rise, and human error can also play a role in choosing the incorrect mode of delivery. To mitigate these risks, a machine learning-based decision-making model has been developed to predict the most appropriate mode of childbirth. This will help to reduce the possibility of incorrect decisions and ensure the safety of both mother and baby

**Introduction:**

Machine Learning has transformed every aspect of life, including business and trade, social and electronic media, education and learning, manufacturing industries, medicine and sciences, and every other sector. New reforms and advanced artificial intelligence technologies have enabled data analysts to transform raw data generated by these industries into meaningful insights for an effective decision-making process. Health care is one of the critical sectors in which a large amount of data is generated on a daily basis, making effective decisions based on this data a challenge.

Machine learning is becoming more prevalent and important in the healthcare industry in predicting and identifying quality treatments for patients and improving other healthcare services. As a result, machine learning techniques are used to extract knowledge from massive and complex data sets in an organized format, which can then be used to make effective decisions. According to Sana et al. [1], machine learning approaches offer diagnosis and analysis amenities in many medical domains and their applications in clinical variables and analytics, such as disease prediction, decision-making based on extracted medical knowledge, and patient management. Furthermore, as the amount of available data increases, machine learning techniques have significant benefits as prediction tools in health care, providing sometimes surprising prediction models that aid in clinical counseling. These tools are essential in biomedical research and are used as part of the clinical decision-making process.

In hospitals, several methods of child delivery are available, but the most common are traditional vaginal birth or cesarean (c-section), with vacuum extractions and obstetric pincers used during complications in vaginal deliveries. There are several assumptions regarding the mode of delivery, but accurately predicting the type of childbirth remains difficult.

**II Literature Review:**

Khan et al., 2020 [2] Conducted a study on Predicting Cesarean Childbirth using Ensemble Machine Learning Methods. The goal of this study is to use data mining to predict whether or not a cesarean section is necessary, thereby increasing the safety of the mother and newborn during and after childbirth by avoiding unnecessary cesarean sections. Three different ensemble prediction models based on XGBoost, AdaBoost, and Catboost were developed to achieve the goal. As a result, XGBoost had the highest accuracy-88.91%, AdaBoost had 88.69% accuracy, and Catboost had 87.66% accuracy.

Kamat et al., 2021 [3] and colleagues conducted research on data generated during the nine months of pregnancy. They use various techniques such as cleaning, sorting, and classification to make the best predictions of anomalies and threats. They primarily considered several parameters, including age, body mass index, parity, glucose fasting, and so on…They created two algorithms using decision trees and naive Bayes based on these parameters. The above parameters are used to predict the curbing cases as the cases increase during delivery to determine the best section to take. The numeric ranges were converted into categorical values after the parameters were selected. Using the ratio of 0.67, the data were divided into training and testing segments. This existing system only takes into account predefined variables and ignores all other parameters that could be used for prediction. Additionally, if one single value is missed, the system will not proceed.

Kowsher [4] and Colleagues proposed a computerized method of decision-making for selecting the appropriate mode of childbirth. Since the process is computerized, the machine learning algorithms are the best because of less error rate. They have used many different classifiers to make the decision more accurate and real. The dataset they considered has many null and null values, so to fill those, they used KNearest neighbors. The considered dataset has 3 categorical variables including presentation, reason, and membrane. Since the values are categorical they have been encoded into numbers by using one-hot encoding. Finally, they perform analysis on different algorithms using 11 statistical measurements like accuracy, FI score, PS, RS, FBS, HL, and others. They compared the performance of all the proposed models among them, the decision tree classifier predicted most with an accuracy of 91.8%.

Zahid Ullah, and Farukh Saleem [5] researched on normal and Cesarean delivery. They first examine the existing data and also previous medical records about the mode of delivery were investigated using machine learning algorithms. They extracted some meaningful insights from unseen cases. So Various prediction Models like decision tree, AdaBoostM1, bagging, and k-nearest neighbors were used to train the model. In this first, the existing data was enriched because this it would increase the result of accuracy in order to predict the mode of childbirth

Muhammad Nazrul Islam, Tahasin Mahmud [6], and others explore the best possible features for predicting the mode and by using these features they have proposed five different algorithms. They conducted interviews and structured reviews to identify the best possible features. Finally, they revealed that 32 features are the best suitable prediction mode of childbirth. These features are grouped and proposed algorithms are applied for analysis and prediction. Based on the Evaluation parameters they finally concluded that the stacking classifier(FI score is 97.4%) is the best.

**III Proposed Work:**

A maternal healthcare application that assists doctors in predicting a mode of delivery compatible with the mother's characteristics would be beneficial in reducing childbirth complications. Various Machine Learning algorithms, such as K-nearest neighbors, Random Forest, and SVM, will be used in this. We will predict which mode of delivery is appropriate at that time based on the accuracy of these algorithms. We use Naive Bayes, KNN, SVM, Random Forest, and XGBoost.

**Dataset:**

The data set contains 7 features, namely, age, Height, Weight,delivery\_number, delivery\_time, blood\_of\_pressure, and heart\_problem, while Mode Of Childbirth, is a class attribute to label whether the delivery was performed by cesarean, Normal, or Emergency.

**Prediction Models:**

**Naïve Bayes:**

Naive Bayes is a machine learning algorithm that is based on the Bayes Theorem. In Naive Bayes, the posterior probability is used to predict the class based on attributes. It is only used in a wide range of classification tasks.

**K-Nearest Neighbour:**

K-Nearest Neighbour algorithm, also known as the KNN algorithm, is a machine learning algorithm that is based on the supervised learning model. The KNN algorithm works by assuming that similar things exist nearby. KNN is a highly adaptable algorithm that can be used to solve a wide range of problems.

**SVM:**

Support Vector Machine, or SVM, is a classification and regression algorithm. However, it is primarily used in Machine Learning for Classification problems. The SVM algorithm's main goal is to find the best line or decision boundary that can divide n-dimensional space into classes so that we can easily place new data points in the right category in the future.

**Random Forest:**

Random Forest can be used in ML to solve both classification and regression problems. It is based on the concept of ensemble learning, which is a process that involves combining multiple classifiers to solve a complex problem and improve the model's performance.

**XGBoost:**

XGBoost is a distributed gradient boosting library that has been optimised for efficient and scalable training of machine learning models. It's an ensemble learning method that combines the predictions of several weak models to produce a more accurate prediction. Because of its ability to handle large datasets and achieve state-of-the-art performance in many machine learning tasks such as classification and regression, XGBoost has become one of the most popular and widely used machine learning algorithms.