IJIRCCE

e-ISSN: 2320-9801 | p-ISSN: 2320-9798

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

Volume 11, Issue 7, July 2023

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

Impact Factor: 8.379

International Journal of Innovative Research in Computer and Communication Engineering

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 |

|| Volume 11, Issue 7, July 2023 ||

| DOI: 10.15680/IJIRCCE.2023.1107032 |

Machine Learning For Pregnancy Risk Prediction

Bhoomika K G[1], Adarsh M J[2]

PG Student, Dept. of Master of Computer Applications, Jawaharlal Nehru New College of Engineering, Shivamogga, India

Assistant Professor, Dept. of Master of Computer Applications, Jawaharlal Nehru New College of Engineering, Shivamogga, India

ABSTRACT: Women are more at-risk during pregnancy, necessitating more medical care and attention. Complications of pregnancy, which result from modifications in physiological parameters, can seriously harm the mother's health. The C4.5 decision tree classification algorithm and the Naive Bayes Classification Algorithm are the two classification algorithms used in this study to predict probable health problems for expectant women. These algorithms make use of data collected from expectant mothers at various phases of their pregnancies to forecast their health and spot possible issues. The study attempts to evaluate a person's risk during pregnancy and lower maternal and foetal mortality by contrasting various methods.

KEYWORDS: Machine learning, C4.5 decision tree classification algorithm, Naive Bayes Classification Algorithm

I. INRODUCTION

As a result of hormonal changes, blood pressure, temperature, weight, glucose levels, infections, and other factors, pregnancy is a delicate period for a woman's body. It might be challenging to return a pregnant woman's health to normality in emergency conditions caused by unqualified medical professionals and pregnant people. According to a recent study, complications during pregnancy and delivery are to blame for about 800 fatalities per day, with underdeveloped countries accounting for the majority of these deaths. 2,890 pregnant women lost their lives in 2013 while carrying children or giving birth, and the majority of these tragedies might have been avoided. In impoverished countries, pregnant women may not always have access to healthcare, which makes pregnancy irresponsibility disastrous. By recognising physiological factors like blood pressure, blood glucose level, and weight that may create problems, one might prevent or reduce these causes of death. Popular data-analysis techniques like the C4.5 decision tree algorithm and naive bayes algorithms can be used to forecast the health of pregnant women. The paper summarises the findings of several analytical techniques and assesses the most effective algorithm for determining the health of pregnant women.

II. RELATED WORK

Kalpesh Adhatrao et al [1], A method developed by Kalpesh Adhatrao and his colleagues makes use of data mining techniques to forecast a student's future performance based on their present performance. They used the ID3 and C4.5 classification algorithms to examine the student data and predict how each new freshman will do on forthcoming examinations.

Mohammed M Mazid [2], When Mohammed M. Mazid and colleagues' suggested technique of picking the most crucial qualities from a dataset is used, the C4.5 algorithm may be made more effective and produce better results.

Tina R. Patil et al [3], this research has shed light on performance evaluation using naive bayes and J48 (C4.5), which is based on successful and unsuccessful instances of data categorization. Instead of focusing on improving their accuracy, they have chosen to assess both classifiers in light of the bank dataset. They can do this to reduce the false positive rate to a minimum while increasing the real positive rate of identifying defaulters.

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 |

|| Volume 11, Issue 7, July 2023 ||

| DOI: 10.15680/IJIRCCE.2023.1107032 |

Li Li et al [4], Both women and their unborn children face a life-changing event during pregnancy, with many of them being more susceptible because of things like early care and routine checkups. PregnancyLine, a visual analytics tool that the researchers created, offers a care plan and explains medical testing at different stages of pregnancy. Increasing medical staff knowledge of patients with anomalies and fostering greater communication between expectant mothers and their doctors are the objectives. The effectiveness of this strategy in educating pregnant women about pregnancy risks, the importance of exams, and recognising anomalies is demonstrated by evaluation findings and case studies.

Erly Krisnanik et al [5], Due to the high expense of therapy, the lack of understanding about pregnancy symptoms, and periodic checks, pregnant women face substantial risks throughout pregnancy. A Pregnancy Risk Detections System (PRDS) is being created to measure a woman's pregnancy risk level based on symptoms in order to resolve this problem. The system is divided into four categories: illness progression, problematic obstetric history, problematic maternal condition, and pregnancy problems. To create advice for midwives, the system employs an analytical process that evolves from descriptive to predictive and prescriptive. With the use of this technique, midwives can more accurately determine the risk of their patients and offer preventative treatment to lower maternal and foetal fatalities.

III. PROPOSED METHODOLOGY AND DISCUSSION

1. Data collection:

By examining previous pregnancy data, this study evaluates pregnancy risks using the C4.5 and Naive Bayes classification algorithms. The collection of raw data, its preparation and organisation, analysis, and assessment of the performance of the created models are the four stages.

2. Preprocessing of data:

In order to ensure correctness, completeness, and usefulness, data cleansing entails analysing databases for anomalies, mistakes, missing information, and unnecessary data. The dataset is separated into training and test sets when normalisation and data splitting are completed.

3. Data splitting:

Data distribution is random; a decision tree model was created, and its accuracy was evaluated using 230 test samples and 370 training samples.

4. Algorithm:

With the use of C4.5 decision trees, pregnant data, and naive bayes, research tries to create prediction-based learning models.

5. Accuracy:

To evaluate the precision of prediction and classification, learning models are subjected to analysis and interpretation using test data. The most trustworthy C4.5 and Naive Bayes algorithms for estimating the risk of pregnancy are chosen using data.

IV. DESIGN AND IMPLEMENTATION

This study will examine and forecast the risks connected to various pregnancies using a variety of categorization information. Out of all the algorithms used in the Decision Tree Induction technique, the C4.5 Algorithm was chosen since it is one of the most effective, potent, and well-known algorithms for categorising medical data. The Naive Bayes Algorithm is one of the robust and successful Bayesian Network methods. The task of classifying medical records is well suited to this algorithm.

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 |

|| Volume 11, Issue 7, July 2023 ||

| DOI: 10.15680/IJIRCCE.2023.1107032 |

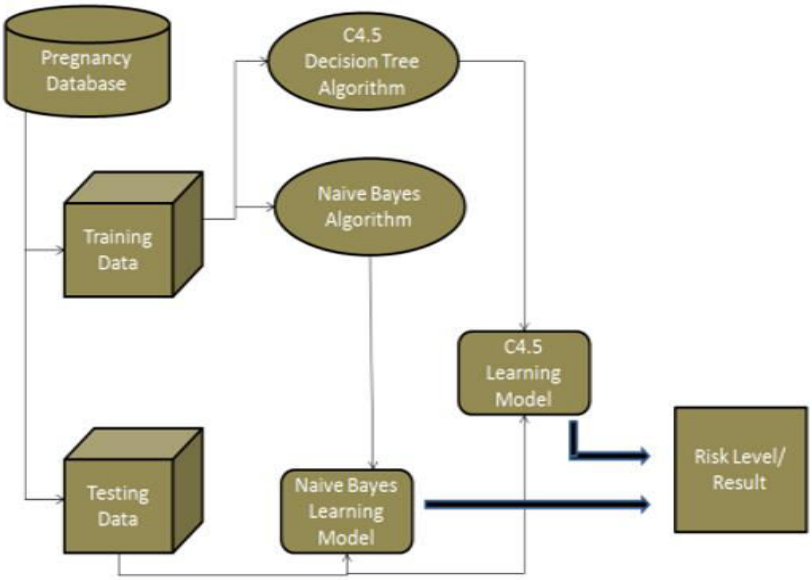

Figure 1: System Architecture of Project

C4.5 Decision Tree Algorithm: Entropy and information gain are used by the decision tree method C4.5 to divide data into more homogenous groupings. It can deal with missing data and qualities that are both continuous and discontinuous. It performs trimming after building the first tree to avoid overfitting. Due to its efficacy and adaptability, C4.5 is frequently used for classification and regression applications.

Naive Bayes Algorithm: A well-liked and effective supervised machine learning approach for classification problems is naive bayes. The Bayes theorem, which determines the conditional probability of an occurrence based on knowledge of related events, is the basis of the algorithm. It works with labelled datasets in the context of Naive Bayes, where each occurrence is represented by a collection of characteristics and an associated class label. Naive Bayes determines the prior probabilities of each class and the chance of each characteristic occuring for each class during the training phase.

The fundamental tenet of Naive Bayes is that the features are conditionally independent given the class label, which means that while the class label is known, the presence or lack of one feature does not impact the presence or absence of another. The technique employs the Bayes theorem to determine the posterior probability of each class given the observed attributes in order to categorise a new instance. The projected class label for the new instance is changed to reflect the class with the highest posterior probability. Gaussian Naive Bayes, Multinomial Naive Bayes, and Bernoulli Naive Bayes are a few variations of Naive Bayes, each of which is best suited for a certain kind of data and feature. Although Naive Bayes is computationally efficient and performs well with high-dimensional data, it may falter when dealing with unbalanced datasets or situations where the independence condition is violated. However, it can serve as a rapid and practical baseline technique for a variety of classification problems.

International Journal of Innovative Research in Computer and Communication Engineering

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 |

|| Volume 11, Issue 7, July 2023 ||

| DOI: 10.15680/IJIRCCE.2023.1107032 |

V. RESULTS AND DISCUSSIONS

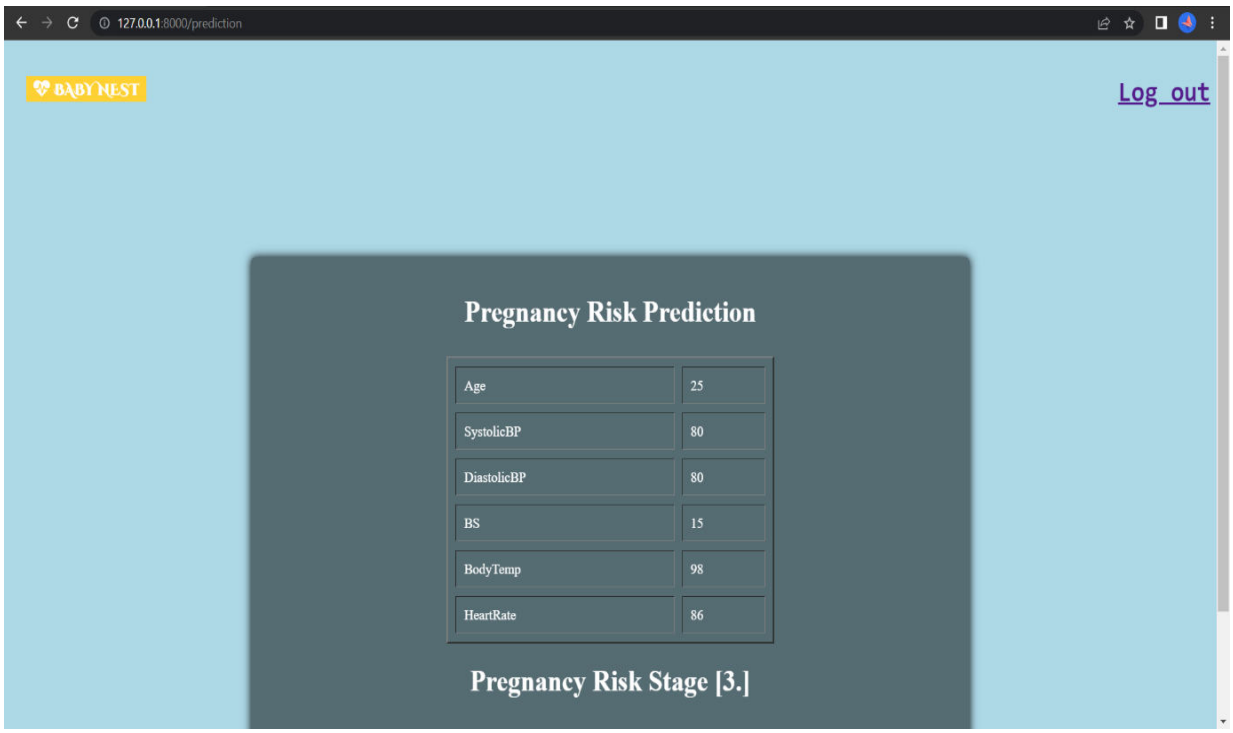

Figure 1 : Pregnancy Risk stage Prediction page

The above Figure 2 shows the Pregnancy Risk Stage using Age, SystolicBP, DiastolicBP, BloodSugar, BodyTemprature and Heart Rate using C4.5 Decision Tree Algorithm and Naive Bayes Algorithm.

VI. CONCLUSION

The C4.5 and Naive Bayes algorithms are effective and strong for classifying and forecasting risks associated with pregnancy. Pregnancy data mining and other types of data mining frequently employ these techniques. The C4.5 decision tree classifier gives superior accuracy in predicting risk levels, according to the study's comparison of the accuracy of C4.5 and Naive Bayes classifiers. The C4.5 classifier outperforms the Naive Bayes classifier when both techniques are applied on 230 data samples. This study emphasises the advantages of prediction-based health monitoring of pregnant women using the C4.5 decision tree classification algorithm.

REFERENCES

1. Predicting Students' Performance Using Id3 And C4.5 Classification Algorithms by Kalpesh Adhatrao, Aditya Gaykar, Amiraj Dhawan, Rohit Jha and Vipul Honrao, Department of Computer Engineering, Fr. C.R.I.T., Navi Mumbai, Maharashtra, India in International Journal of Data Mining & Knowledge Management Process (IJDKP) Vol.3, No.5, September 2013, DOI : 10.5121/ijdkp.2013.3504.
2. Improved C4.5 Algorithm For Rule Based Classification by Mohammed M Mazid, A B M Shawkat Ali, Kevin S Tickle, School Of Computing Science, Central Queensland University, Australia in Recent Advances in Artificial Intelligence, Knowledge Engineering And Data Bases, ISSN: 1790-5109, ISBN: 978-960-474-154-0.
3. Performance Analysis of Naive Bayes and J48 Classification Algorithm for Data Classification by Tina R. Patil, Mrs. S. S. Sherekar, Sant Gadgebaba Amravati University, Amravati, tnpatil2@gmail.com, ss_sherekar@rediffmail.com in International Journal Of Computer Science And Applications, Vol. 6, No.2, Apr 2013, ISSN: 0974- 1011 (Open Access) NCAICN-2013, PRMITR,Badnera.

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 |

|| Volume 11, Issue 7, July 2023 ||

| DOI: 10.15680/IJIRCCE.2023.1107032 |

4. PregnancyLine: A Visual Analysis Method for Maternal Health and Risk Communication by Li Li, Zhigui Liu, Xiao Wang, Xiaoqian Zhang September 2022.
5. Analysing Expert System Data for Research on Risk Factors in Women Who Are Already Pregnant by Erly Krisnanik; Kraugusteeliana Tambunan October 2019.

INNO SPACE
SJIF Scientific Journal Impact Factor

Impact Factor: 8.379

doi cross ref

ISSN INTERNATIONAL STANDARD SERIAL NUMBER INDIA

निस्केयर NISCAIR

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

9940 572 462 6381 907 438 ijircce@gmail.com

www.ijircce.com

Scan to save the contact details