RESEARCH ARTICLE

A Novel Active Learning Technique for Fetal Health Classification Based on XGBoost Classifier

KAUSHAL BHARDWAJ [ID], (Member, IEEE), **NIYATI GOYAL** [ID], **BHAVIKA MITTAL,**
VANDNA SHARMA [ID], **AND SHIV NARESH SHIVHARE** [ID], (Senior Member, IEEE)

School of Computer Science Engineering and Technology, Bennett University, Greater Noida, Uttar Pradesh 201310, India

Corresponding author: Shiv Naresh Shivhare (shiv827@gmail.com)

ABSTRACT Ensuring safe pregnancy and reducing maternal and infant mortality rates require early prediction of fetal health. The application of machine learning algorithms in monitoring fetal health helps to improve the chances of timely intervention and better outcomes in the event of any possible health issues in fetuses. Existing studies offered to help this issue, typically by training models using a significant portion of the dataset, ranging mainly above 70%. The only existing active learning method in this field employs around 41% training samples to achieve 98% accuracy. This work presents a novel active learning technique to identify the most informative data samples to train a model, leading to high accuracy with a limited number of training samples. It employs a novel query function built upon uncertainty and diversity criteria which are derived based on properties of XGBoost classifier and distance from each other. For deriving uncertainty criterion the soft probabilities obtained for the unlabeled samples are used, while the distance among the uncertain samples in feature space is utilized for deriving diversity criterion. The proposed approach shows superior performance compared to all state-of-the-art methods. Through analysis and experimentation, the proposed solution achieves an accuracy greater than 99% using less than 20% of the dataset for training. This shows its efficacy and potential in the monitoring of fetal health. The code and dataset are available on the GitHub repository https://github.com/niyg7/fetal-health-dataset.

INDEX TERMS Fetal health, active learning, XGBoost, query function, uncertainty, diversity.

I. INTRODUCTION

During pregnancy, a pregnant woman goes through various physiological changes. These changes may affect the health of the fetus inside the womb. There are multiple factors that can affect the health of a fetus, such as maternal health conditions, genetic factors, and uterine abnormalities [1]. There have always been cases of maternal infections, high blood pressure, difficult abortions, maternal hemorrhage, and obstructed labor. In very serious conditions, there have been cases of death of the woman giving birth or death of the infant. To identify the higher risks in pregnancy that can lead to the above-mentioned conditions, a diagnosis of fetal health is necessary.

Cardiotocography (CTG) is widely used for the diagnosis of fetal health. CTG is a medical technique that continuously

records and monitors the fetal heart rate and uterine contractions during pregnancy. This method helps assess fetal well-being, detect possible risks, and identify conditions that could require early intervention. The CTG data provide vital information on fetal responses to uterine activity, making it a widely used tool to diagnose fetal health risks. This identification can help obstetricians design the patient's health plan to reduce the possibility of complications that the patient and fetus may face in the future.

Studies have suggested that if problems related to fetal health are identified in the early stages of pregnancy, complications can be avoided by proper treatment and care from the beginning of pregnancy [2]. Fetal heartbeat and uterine contractions of the mother, and many other such characteristics can be used to predict fetal health at an early stage of pregnancy.

The health status of the fetus can be identified using machine learning algorithms and many models have been

The associate editor coordinating the review of this manuscript and approving it for publication was Vishal Srivastava.

introduced for doctors and patients to check the health status of the fetus [3], [4], [5]. Machine learning plays a crucial role in fetal health classification by analyzing various medical parameters such as fetal heart rate, movement, and uterine contractions. These parameters often require constant monitoring and medical expertise, which can be time-consuming and costly. Machine learning algorithms can help identify patterns and underlying behaviors in patient data, allowing for early detection of potential risks associated with fetal health. This early identification of fetal health conditions enables medical experts to take preventive measures at an early stage that can improve the outcomes for both the fetus and the mother. Different machine learning methods have varying properties and thus produce different outcomes [6], [7]. Ensemble learning models have been commonly used in various studies to classify the state of fetal health [8], [9], [10]. Previous studies have concluded that XGBoost, an ensemble learning model, is robust and effective for the classification of fetal health status [11], [12], [13]. This study uses XGBoost with active learning for the classification of fetal health as normal, suspect, or pathological.

Several studies have highlighted the potential of machine learning in predicting fetal health conditions. Akbulut et al. [14] developed an e-health application using maternal clinical history and the Decision Forest model to predict fetal health conditions. Likewise, Yin and Bingi [15] used oversampling and Support Vector Machine (SVM) for predicting fetal health using CTG data. Mehbodniya et al. [16] implemented multiple solutions using SVM, Random Forest (RF), Multi-layer Perceptron (MLP), and K-Nearest Neighbors (KNN) algorithms and highest accuracy was achieved with RF. In addition, Kuzu and Santur [17] illustrated the effectiveness of a predictive technique based on ensemble learning to classify fetal health during non-stress tests (NST). Abiyev et al. [18] introduced a Type-2 fuzzy neural system (T2-FNN) that achieves remarkable precision using CTG data. Hoodbhoy et al. [19] used the XGBoost model to predict abnormal fetal states using machine learning in CTG data. Sahin and Subasi [20] used eight machine learning algorithms on different CTG data with 1831 instances and 21 features and achieved maximum accuracy with RF.

In most of the existing methods, the authors have used 70% to 80% labeled data to train machine learning model models such as SVM, RF, and XGBoost. However, fetal health classification must deal with the challenge of limited labeled samples. Since the labeling of samples has to be done by experts, it is costly and time-consuming. For such a problem, where the availability of training samples is limited (less), we need to choose informative training samples. Active Learning (AL) is a method that helps identify informative samples and provides accurate results with fewer samples. AL is a machine learning approach that optimizes the performance of the model utilizing the minimum possible samples that hold the most crucial information [21], [22], [23], [24], [25]. In AL, we start with a very small number of labeled samples, and in each iteration we select the

informative unlabeled samples, provide them labels, and add to the training dataset. This is performed iteratively until the stopping criteria are met. Therefore, for the classification of fetal health, we employ the AL approach in the XGBoost classifier and iteratively train our model with the informative training samples it selects in each iteration.

To our knowledge, only one AL method [26] has been proposed in the literature for the classification of fetal health. It is based on the query-by-committee technique which was able to generate an accuracy of around 98% using 41% of the data as a training dataset. Further, in each iteration of selection it uses genetic algorithm-based optimization which is time consuming.

Here, we propose a novel AL technique that involves uncertainty and diversity criteria using the properties of the XGBoost classifier. For uncertainty of a sample, its probability of belonging to different classes is predicted using the steps of the XGBoost classifier, and the difference between two maximum probabilities is considered. Thus, n uncertain samples are selected with a minimum difference value. Among these n samples k diverse samples are chosen by considering the distance among them in feature space. The process is repeated until the stopping criteria are met.

The contribution of the proposed method for accurate fetal health classification using a small number of training samples is threefold:

1) An active learning technique with a novel query function based on uncertainty and diversity criteria is proposed to select informative samples to train the classifier model.
2) The properties of the XGBoost classifier are used to determine informative samples.
3) An average accuracy of 99.38% is achieved by using only 19.75% samples of dataset for training the classifier.

The remainder of this paper is structured as follows. Section II provides a review of the literature that includes a review of relevant work done on fetal health classification. The proposed AL technique with XGBoost is described in Section III. Section IV presents the experimental analysis. Finally, Section V draws the conclusion of this paper.

II. LITERATURE REVIEW

Fetal health refers to the well being of the fetal during pregnancy, whether the health of the fetus is normal, suspect, or pathological. It is important to check fetal health before birth to avoid postpartum consequences. Performing a timely check of the health of the fetus can help doctors reduce the risks of potential complications. Failure to monitor fetal health can lead to many severe outcomes, such as premature delivery, congenital abnormalities, birth defects, etc. Using ML models can help doctors by analyzing historical data to identify issues in advance.

Machine learning-based fetal health prediction has witnessed notable advancements, demonstrating the integration of medical science and technological innovation. Multiple
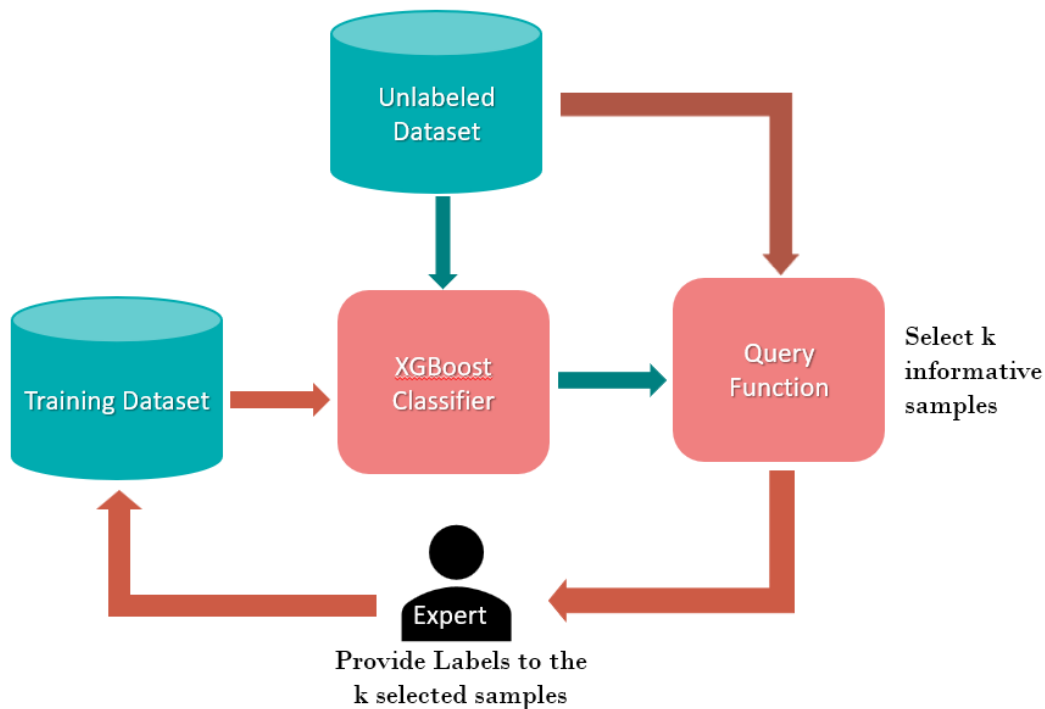
FIGURE 1. General mechanism of active learning. Starts with very few samples in training dataset. In each iteration *k* informative unlabeled samples are selected, labeled by expert and added to the training dataset.

binary classification models such as Averaged Perceptron, Boosted Decision Tree, Bayes Point Machine, Decision Forest, Locally Deep SVM, Logistic Regression (LR), Neural Network and SVM were used to predict fetal health in Akbulut et al. [14]. These models were trained in the clinical dataset of 96 pregnant women. Among the different models, the highest accuracy was demonstrated using the Decision Forest Model when tested on 16 users in real life and the performance was 87.5% which maybe helpful for early fetal health diagnosis. Similarly, in Hoodbhoy et al. [19] different machine learning classification models including Multilayer Perceptron (MLP), XGBoost classifier, Decision tree (DT), RF, LR, SVM linear model, SVM RBF model, KNN, Naive Bayes and AdaBoost were used for prediction of fetal health. Since the training data are imbalanced, the study used the Synthetic Minority Oversampling Technique (SMOTE) method to oversampling the training data. SMOTE generates synthetic examples for the minority class by interpolating between existing instances of that class. The models trained with these samples were then tested with real data. Using experiments on different models, the study found that five main factors have the highest weight in predicting the abnormal fetal state. These factors were short-term variability, percentage of time with abnormal long-term variability, number of accelerations per second (AC), mean value of short-term variability, and uterine contractions. This method achieved an accuracy of 93% using XGBoost.

In another attempt Mehbodniya et al. [16] used multiple ML techniques with 70% training data for fetal health prediction. They also performed a correlation analysis that can help doctors focus on vital parameters. The work also experimented with supervised feature selection techniques, including a filter-based selection method that applied a statistical approach to achieve correlation coefficients between input variables. Ten-fold cross-validation was applied to determine the best parameters while training and achieved the highest accuracy of 94.5% using RF. In Abiyev et al. [18] an integration of type-2 fuzzy sets and neural networks was used to determine the state of fetal health. The construction of the system is based on a cross-validation technique with the gradient algorithm. Employing different numbers of rules, the study found that the increase in the number of rules leads to an increase in the performance of the system. The final accuracy achieved was 96.6%. In addition, an algorithm based on Randomized Input Sampling for the Explanation of Black-Box Models (RISE) was created and named Feature Alteration of the explanation of Black-Box Models (FAB) by Yin and Bingi [15]. Using the FAB algorithm and SVM [27] as their model, it bagged an accuracy of 99.59%.

Furthermore, Sahin and Subasi [20] employed eight different machine learning algorithms that were ANN, SVM, KNN, RF, CART, LR, C4.5 and Radial Basis Function Network (RBFN) in different CTG data which consisted of 1831 instances and 21 attributes. Among the different ML models, KNN gave an accuracy of 98.4% and RF was able to

achieve an accuracy of 99.18%. Kuzu and Santur [17] made use of ensemble learning to achieve an accuracy of 100% using the XGBoost Classifier. They made use of bagging, boosting, and stacking to achieve this accuracy. However, the training set consists of 75% of data. Some other studies have also been conducted related to the classification of fetal health [28], [29], [30], [31], [32].

All of the above-mentioned methods use data 70% or higher for training to achieve precision. As mentioned earlier, the availability of training samples is limited in the case of fetal health classification. When training samples are limited, active learning techniques are useful. To our knowledge, only one study is proposed in the literature based on active learning [26]. The AL method proposed in this work is based on a query-by-committee employing k-means and genetic algorithm. This method integrates the k-means clustering and genetic algorithms based on weighted query-by-committee (KGA-WQBC) and achieved 98.02% accuracy with approximately 41% of training data. Since the training data are costly, 41% training samples is a huge reduction from the other methods; however, further reduction remains a challenge.

III. PROPOSED WORK
In this work, a novel AL technique for fetal health classification based on uncertainty and diversity criteria is presented. The proposed work takes advantage of the properties of the XGBoost classifier in the presented query function to determine the uncertainty of each sample. In addition, the distance between the uncertain samples in the feature space is considered to determine the diverse samples.

AL is a machine learning approach that optimizes the model's performance by utilizing the minimum possible samples that hold the most crucial information. AL is an effective approach when we have a large unlabeled dataset and a very small amount of labeled dataset. In case of fetal health data, the labeling of the unlabeled data samples requires the assistance of medical experts, which is costly and time consuming. Hence, for this purpose an AL approach is suitable to train a model with a very small training dataset. The total dataset U is divided into two sets, M and $U' = U - M$, where the set M is initialized with a m samples ($m = 30$ in the proposed work). AL updates the training dataset M by selecting the most informative samples in each iteration. The module of AL responsible for the selection of the informative samples is called the query function. In each iteration, a bunch of k informative samples is selected from U' and added to M after assigning their class labels. Then, the selected samples are removed from U'. With the updated set M, a new model is trained in each iteration. This process is repeated until the stopping criterion is met.

A. XGBoost CLASSIFIER AND SOFT VOTING PROBABILITIES
The proposed query function of the AL method utilizes the properties of the XGBoost classifier to determine the

uncertain samples. XGBoost classifier is an ensemble method and a boosting technique. The ensemble method combines multiple models to enhance its predictive ability and the boosting technique combines several weak learners to create a strong learner, with each successive learner targeting the missclassifications of its predecessor. In the XGBoost algorithm, multiple decision trees (DTs) are built in a sequential manner to form an ensemble of trees. All the independent variables fed to the DT are assigned some Weights. Before feeding these to the next DT the weights of the wrongly predicted variables are increased. The classifier begins by building a base model that outputs the average of all possible classes as the final prediction for each data sample. Then, it computes the residual for each data sample. Residual, also known as error, is the difference between the actual value and the predicted value. In addition, the classifier builds a new learner, that is, a DT, that works on the residuals of the base model.

The classifier makes a binary split for each level in each tree it builds. For example, for a feature y with three values $y1$, $y2$, and $y3$, it performs a binary split for $y = y1$ and $y = \{y2, y3\}$. The residuals from the base model are considered for the associated data samples in both categories of the split. Since the considered dataset has multiple features, multiple binary splits are possible; thus, to select the best possible feature to split at each level, the 'information gain' is calculated for each split, and then the split with the highest information gain is selected. The information gain can be calculated as shown in Eq.1.

$$\text{Information Gain} = S_L + S_R - S_T \qquad (1)$$

where S_L is the similarity measure of the left side, S_R is the similarity measure of the right side and S_T is the similarity measure of the top branch. The similarity measure can be defined as shown in Eq.2.

$$S = \sum (residual)^2 / P_r + (1 - P_r) \qquad (2)$$

where, P_r is the probability of either the left or right side.

This process is repeated for every new DT it builds, where each new tree acts upon the residuals of the previous tree. After the entire ensemble is constructed, we compute the final predictions for each data sample. Then, a data sample is provided to every tree of the ensemble, and their predictions from each DT are considered as final prediction for that particular data sample.

Another thing to note is that in the proposed work, the XGBoost classifier uses a soft voting mechanism to compute the final probabilities as shown in Fig. 2. Fig 2 shows that each learner gives a different output as class probabilities on which soft voting is performed. To this end, the probabilities obtained from different learners are weighted and the class with the highest averaged probability is selected as the final prediction. In contrast to hard voting, where the class that receives the majority of votes is presented as final prediction outcome, soft voting leverages the results
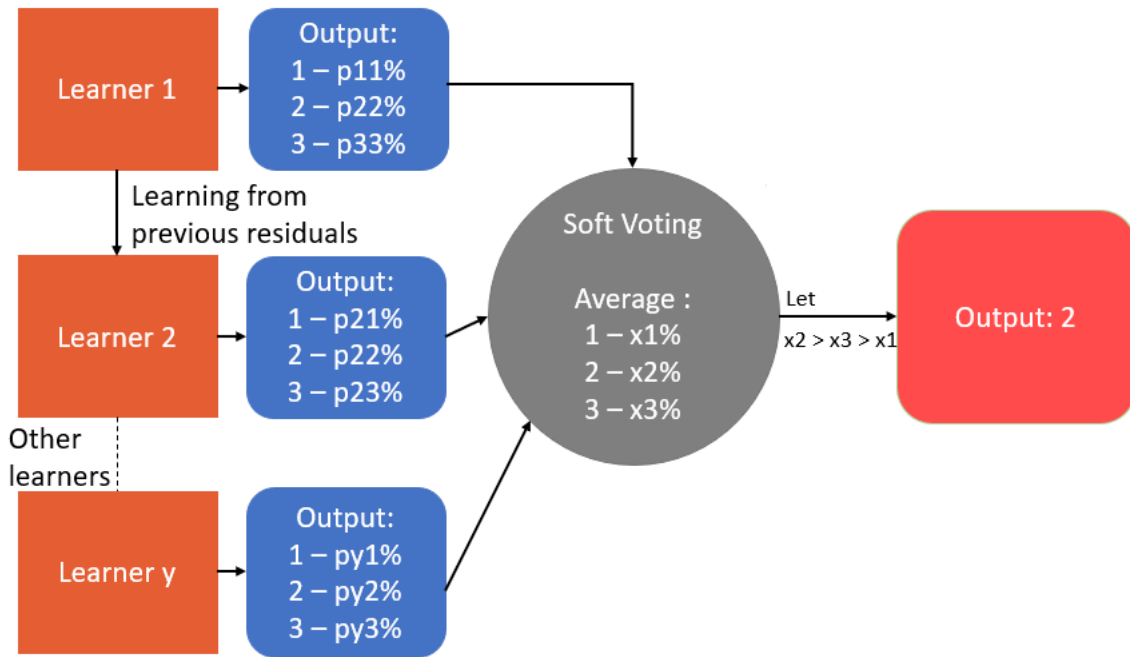
FIGURE 2. Soft Voting Mechanism inside XGBoost. Average of the probabilities from all the learners are considered to determine the final probability for the sample.

of each DT as it provides the weighted average of every tree's predicted probabilities as the predicted output. This process of XGBoost classifier is utilized to obtain the class probabilities for each data sample and these probabilities are used to determine the uncertainty of the samples in the proposed query function.

B. PROPOSED QUERY FUNCTION

The proposed query function attempts to find the most informative samples to train a classifier model to classify the fetal health status. The query function is executed in each iteration of the AL procedure and finds informative samples based on uncertainty and diversity criteria. This procedure is illustrated in Fig. 3. As indicated in the figure, the soft probabilities of every unlabeled sample are taken by the XGBoost classifier trained on available training samples. Subsequently, n uncertain samples are identified on the basis of probabilities. Afterward, k diverse samples are selected from the n uncertain samples. The selected k samples are assigned their class labels and moved from an unlabeled set to a labeled set.

The uncertainty and diversity criteria used in the query function are described as follows:

1) The XGBoost classifier is trained on the training dataset M, which initially consists of m samples.
2) The classifier is provided with the unlabeled dataset U' containing x samples to make predictions.
3) Three Probabilities for each data sample (for belonging to class 0 (normal), class 1 (suspect), and class 2 (pathological)) are computed.
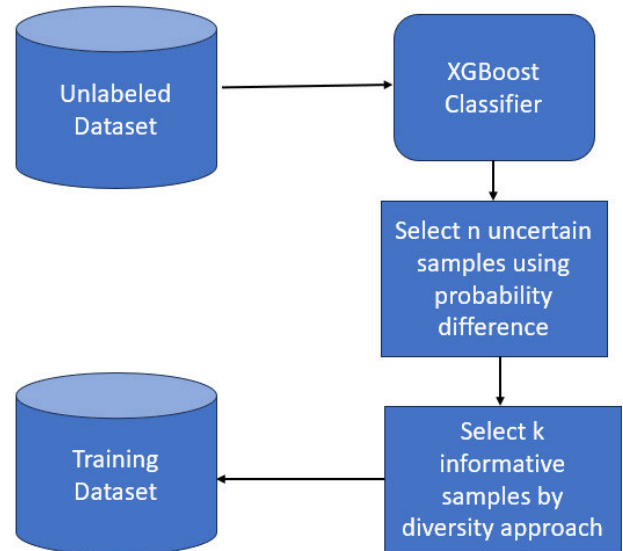

FIGURE 3. The working of proposed query function. For all the unlabeled samples probabilities are obtained by XGBoost classifier. Next n uncertain samples are selected based on uncertainty criterion. Among the n uncertain samples k diverse samples are selected based on diversity criterion and added to the training dataset.

For i^{th} data sample, the probability P^i can be expressed as

$$P^i = \{P_0^i, P_1^i, P_2^i\} \tag{3}$$

4) For each data sample, the difference between the maximum and second maximum probability is calculated and considered as "probability difference".

Algorithm 1 Algorithm of Diversity Criterion

Input: Set of selected samples SP initialized with first uncertain sample, Set of remaining uncertain samples RP having $n - 1$ samples.

Output: set of k selected samples SP.

initialize a variable i to 2

while $i \leq k$ **do**

    initialize $sum\_distances$ to an empty array

    initialize p_index to 0

    while $p\_index < length(RP)$ **do**

        initialize p to $RP[p\_index]$

        initialize $total\_distance$ to 0

        initialize sp_index to 0

        while $sp\_index < length(SP)$ **do**

            $sp \leftarrow SP[sp\_index]$

            $total\_distance \leftarrow total\_distance + distance\_euclidean(SP, p)$ $\triangleright$ finds euclidean distance between samples

            $sp \leftarrow sp + 1$

        end while

        $p\_index = p\_index + 1$

        append $total\_distance$ to the $sum\_distances$ array

    end while

    $farthest\_point\_index \leftarrow argmax(sum\_distances)$ $\triangleright$ Find the index of sample having maximum $sum\_distance$

    append $RP[farthest\_point\_index]$ to SP

    delete $RP[farthest\_point\_index]$ from RP

    $i \leftarrow i + 1$

end while

return SP

For i^{th} sample, the probability difference θ_i is computed as follows

$$max1^i = max\{P_0^i, P_1^i, P_2^i\} \tag{4}$$

$$max2^i = max\{P^i - max1^i\} \tag{5}$$

$$\theta_i = max1^i - max2^i \tag{6}$$

The Samples for which this probability difference is low means the classifier is predicting the two class with similar probabilities and therefore signifies high uncertainty whereas the ones having a significant probability difference represent low uncertainty.

5) We separate n data samples from the unlabeled dataset U' having the least probability difference. These are the samples having the highest uncertainty among the unlabeled dataset.

$$Y = \{\theta^1, \theta^2, \theta^3, \ldots, \theta^n\} \tag{7}$$

Here, θ^1 is the minimum probability difference, θ^2 is the second minimum probability difference, and so on till θ^n, where θ^n is the nth minimum probability difference. The samples corresponding to these probabilities are selected in set S.

$$S = \{S^1, S^2, \ldots, S^n\} \tag{8}$$

S^1 is the data sample from the unlabeled pool having θ^1 probability difference, S^2 is the data sample having θ^2 probability difference and so on.

To pick out the most optimal k data samples from these n data samples, the concept of diversity is applied. It aids us by discovering the most diverse and spread out samples among the n data samples.

6) The data sample with minimum probability difference, i.e., highest uncertainty, is selected and stored in a list, say, $selected\_points(SP)$. The remaining samples are stored in another list, say, $remaining\_points(RP)$.

$$SP = \{S^1\} \tag{9}$$

$$RP = \{S^2, S^3, S^4, \ldots, S^n\} \tag{10}$$

7) Now, as one can see in algorithm 1, the distance of the selected sample is calculated from the remaining $n - 1$ data samples held in the $remaining\_points$.

8) The sample having maximum distance from the selected sample is chosen and appended to $selected\_points$. Subsequently, we remove the data sample from the $remaining\_points$.

9) Now, we recalculate the distance of both the data samples in $selected\_points$ from all points of $remaining\_points$ (i.e., the rest $n - 2$ data samples).

10) For every remaining data sample, distances from both the selected samples are summed. Then, the data sample having the maximum summed distance is selected and appended to $selected\_points$ as well removed from $remaining\_points$.

11) Step 9) and 10) are repeated until there are k data samples in $selected\_points$. These k data samples are the final data samples having the highest information, which we further append to our training dataset for the next dataset.

The training dataset constructed in this way after a certain number of iteration is considered for final training of the XGBoost classifier. The XGBoost classifier trained with these informative samples has high potential in fetal health classification.

C. XGBoost PARAMETER OPTIMIZATION

In order to achieve optimal results, a grid search is performed on the parameters of the XGBoost classifier. For this, a set of parameter values is considered, and experiments are carried out for all combinations. The parameters considered for optimization are the following. Number of estimators, Learning rate, Maximum depth, Minimum child weight, Gamma, Subsample, and Colsample by tree. The parameter values for these parameters are searched in the following range:

- No of estimators: {50, 100, 200}
- Learning rate: {0.01, 0.05, 0.1}
- Maximum depth: {3, 5, 7}
- Minimum child weight: {1, 5}
- Gamma: {0, 0.1, 0.3}

- Subsample: {0.7, 1}
- Colsample by tree: {0.5, 0.8, 1}

After searching the grid, the values of the following parameters are identified as the optimal combination.

- No of estimators: 200
- Learning rate: 0.05
- Maximum depth: 5
- Minimum child weight: 1
- Gamma: 0
- Subsample: 1
- Colsample by tree: 0.5

Note that during the iterations of the active learning algorithm, the XGBoost classifier works only with default values. The grid search is performed after the selection of 420 training samples. Therefore, the XGBoost classifier is re-trained with the estimated optimized parameters and the final training dataset to generate an accurate result.

1) SIGNIFICANCE OF THE CONSIDERED PARAMETERS

1) **No of estimators**: This parameter controls the number of boost rounds. The optimal value for this parameter from the given range is 200. Although more trees improve learning, they also increase training time and, if not regulated, may lead to overfitting.

2) **Learning rate**: This parameter controls the contribution of each tree. Lower values require more boosting rounds, but can improve model accuracy by allowing more gradual learning. The optimal value for this parameter is found to be 0.05, which slows learning but improves performance by increasing the number of boosting rounds.

3) **Maximum depth**: This parameter defines the maximum depth of each tree, which is crucial for controlling the complexity of the model. Deeper trees capture complex patterns with the risk of overfitting, while shallower trees may underfit. The optimized value for this parameter is 5, balancing complexity and the risk of overfitting.

4) **Minimum child weight**: This parameter determines the minimum sum of instance weights required in a child node, effectively controlling the minimum number of instances needed in a leaf. Higher values prevent the model from making overly specific splits, reducing overfitting. The optimal value for this parameter is 1, meaning that only one instance is required for a leaf node.

5) **Gamma**: This parameter defines the minimum loss reduction needed to make a further split. Higher values make the model more conservative, since the splits are discarded if the information gain is below Gamma. In the proposed method, both the initial and optimized values are 0, meaning that a split would occur even with minimal information gain if required.

6) **Subsample**: This parameter controls the fraction of samples used for each boosting round. Using only a subset of data can reduce overfitting but may reduce accuracy if the subset is too small. Here, the initial and optimized values were both 1, meaning 100% of the training dataset was used in each boosting round.

7) **Colsample by tree**: This parameter controls the fraction of features randomly sampled for each tree. During active learning iterations, the default value of 1 was used, which means that all features were considered for every tree. The optimal value for this parameter found according to the grid search is 0.5, indicating that only 50% of the features are considered for each tree. This helps prevent overfitting by using a subset of features for each tree, as a lower value reduces overfitting, although an excessively low value may cause underfitting.

Therefore, the parameter values found after the grid search are used to generate the final results. With the completion of active learning iterations we have highly informative labeled samples and with grid search we have optimal parameter values for training the XGBoost classifier for fetal health classification.

IV. EXPERIMENTAL ANALYSIS
A. DATASET

The proposed method is assessed by conducting the experiments on the CTG dataset collected for fetal health analysis and available both at UCI machine learning repository [33] and the GitHub repository https://github.com/niyg7/fetal-health-dataset. The dataset has 2126 instances, 21 feature variables, and a target variable, i.e., fetal health status. The dataset comprises the measurements of the fetal health rate (FHR) and uterine contraction features on cardiotocograms classified by expert obstetricians. The features capture insights into fetal well-being during pregnancy, e.g., average heart rate of the fetus, temporal increases in fetal heart rate above baseline (average), frequency of fetal movement, contractions of the uterus, and more. All features are integer or continuous values. The target variable is the status of fetal health and has three possible classes: normal, suspect, and pathological. It is a highly imbalanced dataset, comprising of 1655 normal, 295 suspect, and 176 pathological entries in the total 2126 samples. This dataset facilitates analysis of Fetal Heart Rate (FHR) patterns and Uterine Contractions (UC) characteristics to help in obstetric decision-making and monitoring of fetal health. Data supporting the findings of this study are openly available on the GitHub repository https://github.com/niyg7/fetal-health-dataset.

B. EXPERIMENTAL SETUP

To measure the effectiveness of the proposed method, an 8 GB RAM machine and a 64-bit operating system with an x64-based processor are used for the experiment. The machine learning and data analysis tasks are all performed in Python. XGBoost model used for classification has the following objective parameter as ''multi:softprob''.

The value of the initial labeled samples $(m) = 30$, initial unlabeled samples $(x) = 1033$, samples chosen using uncertainty criteria $(n) = 20$, samples chosen using diversity criteria $(k) = 5$.

C. RESULTS

In this section, the results obtained by the proposed method are compared with the state-of-the-art methods for fetal health classification. The best accuracies obtained by the state-of-the-art methods and the average accuracies obtained by the proposed method are reported in Table 1. As shown in the table, Mehbodniya et al. [16] have used RF with 70% training samples to achieve an accuracy of 94.5%. XGBoost model has been trained with 75% of samples to achieve 100% accuracy in Kuzu and Santur [17]. Gradient Boosting was also utilized in two studies, one by Kaliappan et al. [34] that used 80% training samples attaining 99% accuracy, and the other by Nandipati and Xinying [35] that uses 70% training samples attaining 96.78% accuracy. Yin and Bingi [15] illustrated the use of SVM in 75% of training samples, ensuring an accuracy of 99.59%. The stacking ensemble learning technique was used by Bhowmik et al. [36] using the training dataset 75% that reached an accuracy of 96.05%. In Fei et al. [37], the fuzzy C-means clustering technique was utilized in training samples 70%, securing an accuracy of 96.39%. The AL technique presented by Quan et al. [26] that designed the query function based on query-by-committee employing k-means and genetic algorithm and achieved 98% accuracy by using 41% of samples for training.

The proposed work is also compared with two AL techniques designed during experiments. In the first experiment, a batch of 5 most certain samples is selected in each iteration. The certainty of samples is evaluated on the basis of their probability of belonging to a class. In this experiment, the accuracy could not improve above 85% with 420 samples. In second experiment where a batch of 5 samples are selected randomly from the unlabeled set, the accuracy could improve till a level of 95% only with 420 samples used to train.

In the third experiment, which is the proposed method, uncertainty and diversity criteria are applied in the query function as described in Section III. The proposed method achieved an average accuracy of 99.38% with 420 samples. This shows that the proposed method can outperform state-of-the-art methods. This should be noted that, except for AL techniques, other methods in the literature use 70% to 80% data as training to achieve an accuracy above 95%. The existing AL technique uses 41% of the data for training to achieve an accuracy of 98%. However, the proposed method achieves an accuracy of 99.38% with less than 20% training samples. The results of the proposed method can also be visualized in Figure 5. The figure shows that the proposed method could reach the level of accuracy above 97% with 300 samples and 99.38% with 420 samples. At the same time, the AL method that selects certain samples or randomly selects samples cannot reach the level of 95%.
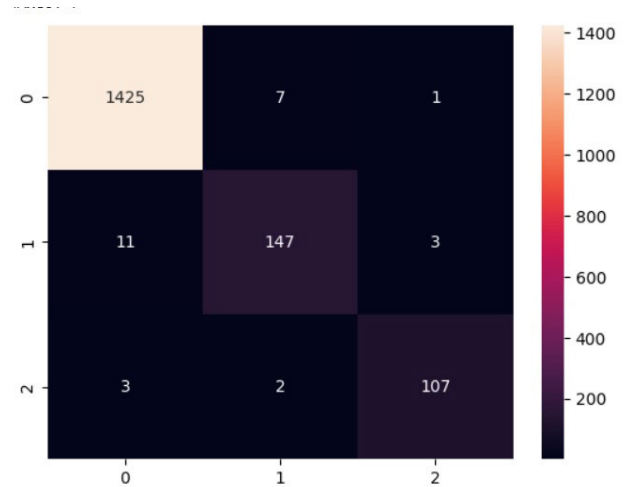

FIGURE 4. The confusion matrix showing the class-wise predictions obtained for the proposed method.
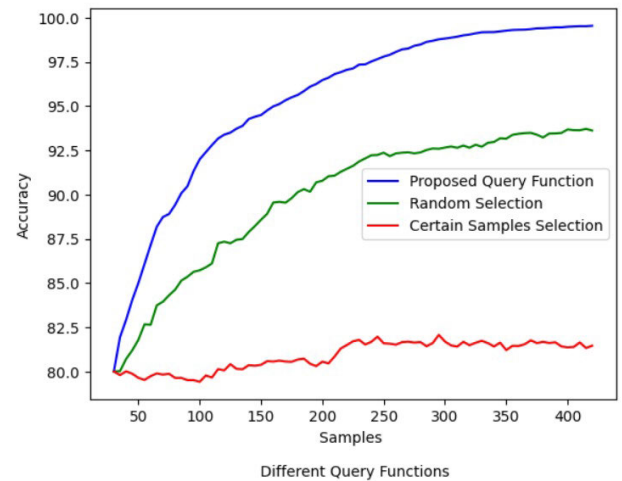

FIGURE 5. Performance of AL method with proposed query functions, Random selection and certain sample selection in fetal health classification.

In order to demonstrate the effectiveness of the proposed method, it is also evaluated for precision, recall, and the F1- score as presented in the table. 2. We can observe that the values obtained are significant for the classification of fetal health. The confusion matrix obtained for the proposed method is shown in Fig. 4. One can also observe from this figure that the proposed method can predict the correct output most of the time. Only a few instances are misclassified due to challenging instances.

D. EFFECT OF INITIAL SAMPLE SELECTION AND SMALL DATASET

A set of experiments is carried out to determine the appropriate selection of the initial samples and assess the performance with unlabeled datasets of different sizes. Two different sets of experiments are carried out with five different sizes of dataset. The first set of experiments

TABLE 1. Performance comparison of proposed method with state-of-the-art methods for fetal health classification.

Author(s)	ML Technique	Training Size	Accuracy
Mehbodniya et al. [16]	RF	1488 (70%)	94.50%
Kuzu and Santur [17]	XGBoost	1594 (75%)	100%
Kaliappan et al. [34]	Gradient Boosting	1700 (80%)	99.00%
Nandipati and Xinying [35]	Gradient Boosting	1488 (70%)	96.98%
Yin and Bingi [15]	SVM	1594 (75%)	99.59%
Bhowmik et al. [36]	Stacking Ensemble Learning	1594 (75%)	96.05%
Fei et al. [37]	Fuzzy C-means clustering	1488 (70%)	96.39%
Quan et al. [26]	AL method based on query by committee employing k-means and Genetic Algorithm	871 (41%)	98.02%
Certain Sample Selection	AL method based on selecting only certain samples	420 (19.75%)	81.61%
Random Selection	AL method based on random selection of informative samples	420 (19.75%)	94.67%
Proposed Method	AL method based on uncertainty and diversity exploiting property of XGBoost classifier	420 (19.75%)	99.38%

(a) Random selection of initial samples   (b) Stratified selection of initial samples
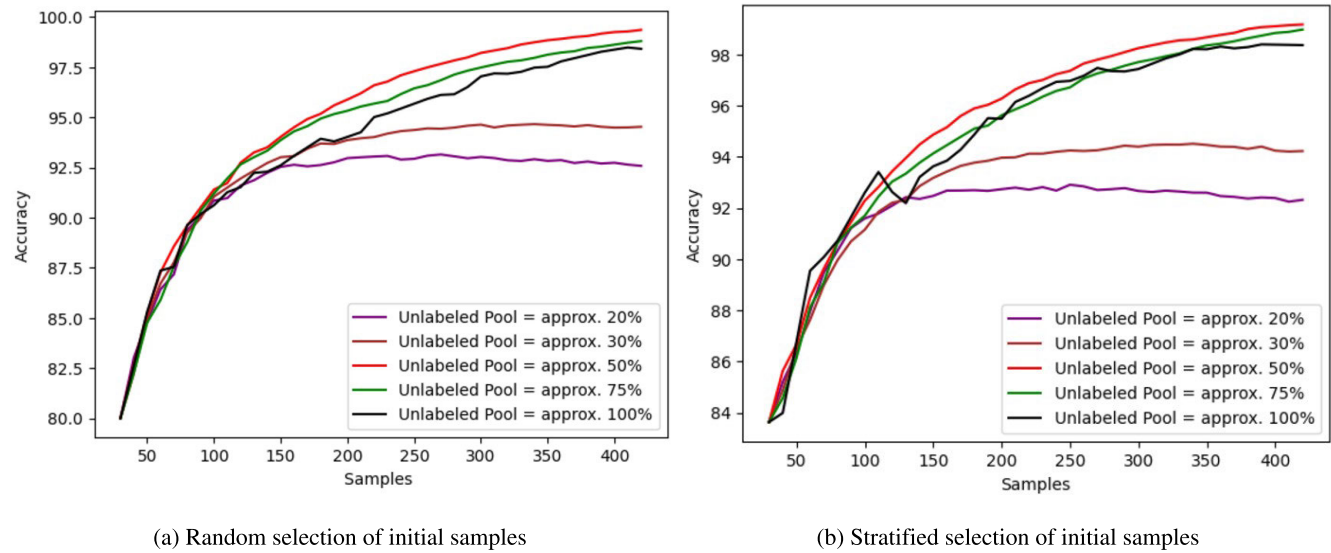
FIGURE 6. Results obtained with different sizes of unlabeled dataset for (a) initial samples selected randomly and (b) stratified selection of initial samples.

TABLE 2. Performance of the proposed method for automated fetal health classification in terms of precision, recall, and F1-score.

Class	Precision	Recall	F1-score
Normal (0)	0.99	0.99	0.99
Suspect (1)	0.94	0.91	0.93
Pathological (2)	0.96	0.96	0.96

is based on the random selection of initial samples, and the second set of experiments is based on the stratified selection of initial samples. In both sets of experiments, five subsets of the dataset are considered namely {20%, 30%, 50%, 75%, 100%}. All experiments start with 30 initial samples.

In the first set of experiments, the initial samples are randomly selected. These experiments were carried out with the five different subsets of the dataset mentioned above. In these experiments, m (= 30) initial samples are selected randomly which are approximately 1.5% and the remaining samples of { 20%, 30%, 50%, 75%, 100%} are considered as unlabeled set. The results obtained from these experiments

are presented in Fig. 6. One can observe from the figure that for initial few iterations, all the experiments are leading to similar results. However, after a few iterations when 90 or more samples are selected, the difference between the accuracies becomes visible in the subsequent iterations. One can clearly see that with around 20% samples the model was able to achieve an accuracy of 92% with 200 samples from which it flattens and with 420 samples it was also able to achieve only 92% accuracy. Similarly with 30% samples, the curve flattens from 200 samples with 93% accuracy. With 50% samples it shows the best performance and achieves the highest accuracy of 99%. When a sample size larger than 50% is considered, the accuracy is almost similar.

In the second set of experiments, the initial samples ($m = 30$) are selected stratified. In other words, the initial samples are selected in such a manner that the samples chosen for each class follow a proportion equal to their class-wise proportion in the entire dataset. The results of this test are presented in Fig. 6b. In these experiments also
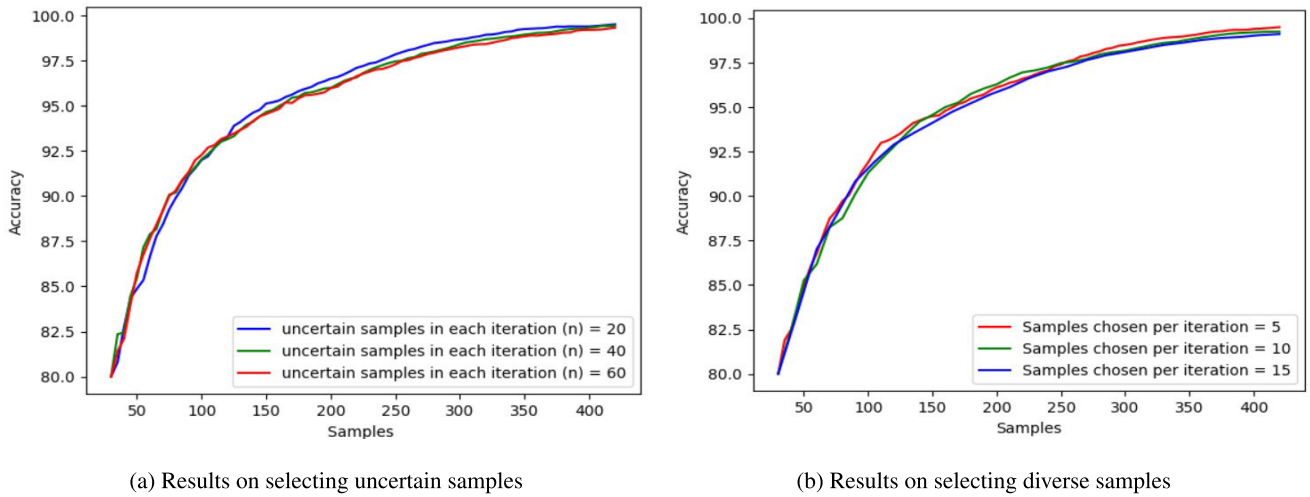
(a) Results on selecting uncertain samples

(b) Results on selecting diverse samples

FIGURE 7. Results obtained for different batch size for (a) uncertain samples *n* varied for 20, 40 and 60 and (b) diverse samples *k* varied for 5, 10 and 15.

the accuracy increases with the size of the unlabeled pool until the size of the unlabeled pool of 50% and after that the accuracy is almost similar or even deteriorates. It is also observed that when the dataset size is 100% the accuracy kept fluctuating more than others. This may be because the method decides the most informative sample in the current iteration based on the XGBoost classifier trained with samples of previous iteration and being a large sample space, it could find less informative samples in some iterations. Although when the dataset size is 50%, the model shows its best performance.

From these two sets of experiments, it is observed that the average accuracy obtained in the case of randomly selected initial samples having 50% unlabeled pool size was 99.22% whereas in the case of stratified initial samples having the same size of 50% for unlabeled pool, the average accuracy obtained was 99%. Therefore, considering the consistent results of the 50% dataset, it is proposed to keep the size of the unlabeled pool of 50% while randomly selecting the initial samples.

E. EFFECT OF BATCH SIZE

Batch size means the number of samples selected in each iteration of the active learning model. In the proposed method, the samples are selected in two stages corresponding to the two criteria. In the first stage *n* uncertain samples are chosen from the unlabeled pool, and in the second step *k* diverse samples are chosen from the *n* uncertain samples. In order to visualize the effect of sample size on the performance of the proposed method, two sets of experiments are carried out. In the first set of experiments, we try to see the effect of the batch size of uncertain samples (*n*) to be selected in the first phase of the query function. In these experiments, the sizes of uncertain samples are varied, or in other words, the value of *n* is varied as $n = 20$, $n = 40$, and $n = 60$. The results obtained are presented in Fig. 7, where one

can observe that although the accuracies attained by the three different values are almost similar, the accuracy achieved by $n = 20$, represented by the blue line, is higher from and after 120 samples. Thus, for further experiments, we choose the value of *n* to be 20.

The second set of experiments are conducted to see the effect of the batch size of diverse samples (*k*) on the performance of the proposed method. These *k* samples are to be selected in the second phase of the query function. Please note that these *k* samples are to be selected from the *n* samples selected in the first phase of the query function. For these experiments, the value of *n* is chosen as 20 based on previous experiments. Three different experiments are carried out to see this effect with the value of *k* as, $k = 5$, $k = 10$, and $k = 15$. The results obtained are presented in Fig. 7b. One can observe from the figure that the accuracies are similar for all three cases. However, when the value of *k* is 5, that is, 5 samples are selected per iteration, the accuracy obtained is slightly better than in other cases. Therefore, in the proposed method it is suggested to select a batch of $n = 20$ uncertain samples in the first phase of the query function and a batch of $k = 5$ diverse samples in the second phase of the query function.

Note that in all the above experiments, it was taken care that the final train size achieved was the same for each of the cases, and that is 420 samples for training. One can see that the best results are achieved when $n = 20$ and $k = 5$ from above experiments.

F. ABLATION STUDY

In order to highlight the importance of the uncertainty and diversity criteria, an ablation study is conducted. To this end, two studies are performed to evaluate the individual effects of the uncertainty and diversity criteria. The experimental setting is kept similar to the proposed method, where the unlabeled pool is 50% and 5 samples are

chosen per iteration, having the final training dataset size of 420 samples.

In the first study, only the uncertainty criterion is considered, where we calculate three probabilities for each sample, P_0 (Probability that the sample belongs to class 0), P_1 (Probability that the sample belongs to class 1), P_2 (Probability that the sample belongs to class 2). Then the difference between maximum and second maximum probability is taken. This probability difference is considered to determine the uncertainty of the sample. The smaller the probability difference, the higher the uncertainty about the true class label of the corresponding sample. Thus, to consider the samples with higher uncertainty, we select 5 samples with smallest probability differences among all the samples in each iteration. The XGBoost model trained on the finally selected training dataset of 420 samples achieves an accuracy of 95.3%.
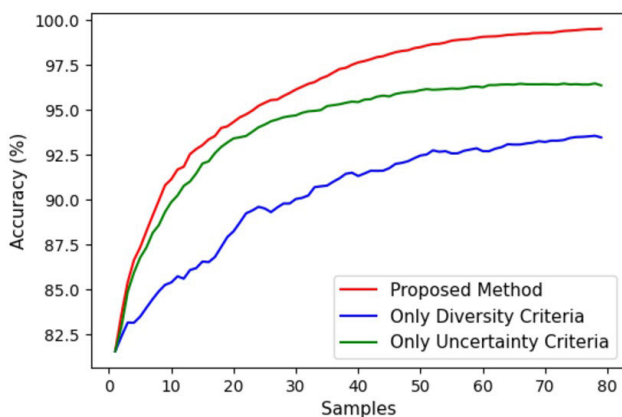

FIGURE 8. Comparison of proposed method which is based on both uncertainty and diversity criteria with methods based on only uncertainty and only diversity.

In the second experiment, only the diversity criterion is considered, where the first sample is randomly selected from the unlabeled pool, and then the distance from the remaining samples is calculated. The sample having the highest distance from the selected sample, that is, the farthest sample in the feature space, is considered and added to the list of selected samples. In the next iteration, the distance of each selected sample from all the remaining samples is calculated again, and then the sum of these distances is taken into account to find the farthest sample. This process is repeated until five samples are selected in each iteration of the unlabeled sample selection. The XGBoost model trained on the selected training dataset of 420 samples achieves an accuracy of 91%.

The proposed model is capable of achieving an accuracy greater than 99% with 420 samples when both the uncertainty and diversity criteria are used. The comparison of the proposed method with the above two experiments is presented in Fig.8. The figure shows that although the uncertainty alone performs similar trends of accuracy in the

initial iterations, after a few iterations the proposed method outperforms all of them. This shows that the individual impact of the uncertainty and diversity criteria is limited in comparison to their combined impact as presented in the proposed method. This also confirms the relevance of combining the uncertainty and diversity criteria in the proposed active learning method for the classification of fetal health.

G. LIMITATION AND FUTURE WORKS
The proposed approach shows huge potential in the classification of fetal health. However, the requirement of 420 training samples is still costly. The future challenge will be to develop effective methods to achieve similar accuracies with fewer training samples. Further the parameters selection for XGBoost classifier is another challenge. In the proposed method, the parameters are chosen after grid search on a limited range of parameter values. Future research may focus on a more extensive search for parameter values for the XGBoost classifier. Furthermore, active learning research in the field of fetal health classification is scarce. Future research may focus on developing novel active learning techniques that utilize XGBoost or other techniques for fetal health classification.

V. CONCLUSION
To ensure safe pregnancies and reduce maternal and infant mortality rates, it is essential to address and understand the various factors that impact fetal health. To this end, multiple studies have been presented suggesting the use of machine learning algorithms in fetal health monitoring to detect early potential health problems in fetuses. All existing solutions given in the literature use around 75%-80% of the dataset for model training. The only active learning technique in the literature requires approximately 40% training samples. Since the labeling of samples requires careful observation by experts and the labeling of the samples is costly, a limited number of training samples are available for the monitoring of fetal health. Therefore, employing large amounts of data for training is extremely expensive. In this paper, a solution to this problem is proposed that utilizes an AL technique to discover the most informative samples to train the machine learning model. The proposed solution uses a novel query function based on uncertainty and diversity criteria. The uncertainty criterion finds the samples for which the current classifier is confused and it is decided based on their soft probabilities obtained by the XGBoost classifier. The diversity criterion chooses the samples which are distant from each other in the feature space. Experiments are carried out on the publicly available data where the proposed solution achieves an average accuracy of 99.38% using less than 20% of the dataset for the training of the XGBoost classifier model and outperforms all of the state-of-art methods. Multiple experiments are carried out to understand the effect of the choice of initial samples and a small dataset on the performance of the proposed method. The best performance

of the proposed method is also supported by an ablation study which shows the joint impact of both the uncertainty and diversity criterion. This confirms the potential of the proposed method for fetal health classification.

DECLARATIONS

The authors declare that there is no conflict of interest with respect to the publication of this article. The authors have no financial or proprietary interest in the material discussed in this article.

REFERENCES

[1] L. Crowe, R. H. Graham, S. C. Robson, and J. Rankin, "A survey of health professionals' views on acceptable gestational age and termination of pregnancy for fetal anomaly," *Eur. J. Med. Genet.*, vol. 61, no. 9, pp. 493–498, Sep. 2018.

[2] S. Belciug and D. G. Iliescu, "Deep learning and Gaussian mixture modelling clustering mix. A new approach for fetal morphology view plane differentiation," *J. Biomed. Informat.*, vol. 143, Jul. 2023, Art. no. 104402.

[3] D. Mennickent, A. Rodríguez, M. C. Opazo, C. A. Riedel, E. Castro, A. Eriz-Salinas, J. Appel-Rubio, C. Aguayo, A. E. Damiano, E. Guzmán-Gutiérrez, and J. Araya, "Machine learning applied in maternal and fetal health: A narrative review focused on pregnancy diseases and complications," *Frontiers Endocrinology*, vol. 14, pp. 1–20, May 2023.

[4] O. C. Olayemi and O. O. Olasehinde, "Machine learning prediction of fetal health status from cardiotocography examination in developing healthcare contexts," *J. Comput. Sci. Res.*, vol. 6, no. 1, pp. 43–53, Mar. 2024.

[5] C. V. Ananth and J. S. Brandt, "Fetal growth and gestational age prediction by machine learning," *Lancet Digit. Health*, vol. 2, no. 7, pp. e336–e337, Jul. 2020.

[6] N. Rahmayanti, H. Pradani, M. Pahlawan, and R. Vinarti, "Comparison of machine learning algorithms to classify fetal health using cardiotocogram data," *Proc. Comput. Sci.*, vol. 197, pp. 162–171, May 2022.

[7] S. Kuhle, B. Maguire, H. Zhang, D. Hamilton, A. C. Allen, K. S. Joseph, and V. M. Allen, "Comparison of logistic regression with machine learning methods for the prediction of fetal growth abnormalities: A retrospective cohort study," *BMC Pregnancy Childbirth*, vol. 18, no. 1, p. 333, Dec. 2018.

[8] M. Al Duhayyim, S. Abbas, A. Al Hejaili, N. Kryvinska, A. Almadhor, and H. Mughal, "Ensemble learning for fetal health classification," *Comput. Syst. Sci. Eng.*, vol. 47, no. 1, pp. 823–842, 2023.

[9] R. Kapila and S. Saleti, "Optimizing fetal health prediction: Ensemble modeling with fusion of feature selection and extraction techniques for cardiotocography data," *Comput. Biol. Chem.*, vol. 107, Dec. 2023, Art. no. 107973.

[10] M. Chidambaram and J. Joy, "Comparative analysis of ensemble learning methods for enhancing fetal health prediction using cardiotocography," *Int. J. Scientific Res. Comput. Sci., Eng. Inf. Technol.*, vol. 2, pp. 217–224, Nov. 2022.

[11] A. Ogunleye and Q.-G. Wang, "XGBoost model for chronic kidney disease diagnosis," *IEEE/ACM Trans. Comput. Biol. Bioinf.*, vol. 17, no. 6, pp. 2131–2140, Nov. 2020.

[12] J. Chen, F. Zhao, Y. Sun, and Y. Yin, "Improved XGBoost model based on genetic algorithm," *Int. J. Comput. Appl. Technol.*, vol. 62, no. 3, p. 240, 2020.

[13] H. Li, Y. Cao, S. Li, J. Zhao, and Y. Sun, "XGBoost model and its application to personal credit evaluation," *IEEE Intell. Syst.*, vol. 35, no. 3, pp. 52–61, May 2020.

[14] A. Akbulut, E. Ertugrul, and V. Topcu, "Fetal health status prediction based on maternal clinical history using machine learning techniques," *Comput. Methods Programs Biomed.*, vol. 163, pp. 87–100, Sep. 2018.

[15] Y. Yin and Y. Bingi, "Using machine learning to classify human fetal health and analyze feature importance," *BioMedInformatics*, vol. 3, no. 2, pp. 280–298, Apr. 2023.

[16] A. Mehbodniya, A. J. P. Lazar, J. Webber, D. K. Sharma, S. Jayagopalan, P. Singh, R. Rajan, S. Pandya, and S. Sengan, "Fetal health classification from cardiotocographic data using machine learning," *Expert Syst.*, vol. 39, no. 6, pp. 1–20, Jul. 2022.

[17] A. Kuzu and Y. Santur, "Early diagnosis and classification of fetal health status from a fetal cardiotocography dataset using ensemble learning," *Diagnostics*, vol. 13, no. 15, p. 2471, Jul. 2023.

[18] R. Abiyev, J. B. Idoko, H. Altıparmak, and M. Tuzunkan, "Fetal health state detection using interval type-2 fuzzy neural networks," *Diagnostics*, vol. 13, no. 10, p. 1690, May 2023.

[19] B. Hasan, Z. Hoodbhoy, M. Noman, A. Shafique, A. Nasim, and D. Chowdhury, "Use of machine learning algorithms for prediction of fetal risk using cardiotocographic data," *Int. J. Appl. Basic Med. Res.*, vol. 9, no. 4, p. 226, 2019.

[20] H. Sahin and A. Subasi, "Classification of the cardiotocogram data for anticipation of fetal risks using machine learning techniques," *Appl. Soft Comput.*, vol. 33, pp. 231–238, Aug. 2015.

[21] K. Bhardwaj, A. Das, and S. Patra, "Spectral–spatial active learning with superpixel profile for classification of hyperspectral images," in *Proc. 6th international conference signal processing*, 2020, pp. 149–155.

[22] S. Patra, K. Bhardwaj, and L. Bruzzone, "A spectral–spatial multicriteria active learning technique for hyperspectral image classification," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 10, no. 12, pp. 5213–5227, Dec. 2017.

[23] S. Budd, E. C. Robinson, and B. Kainz, "A survey on active learning and human-in-the-loop deep learning for medical image analysis," *Med. Image Anal.*, vol. 71, Jul. 2021, Art. no. 102062.

[24] S. K. Sudha and S. Aji, "An active learning method with entropy weighting subspace clustering for remote sensing image retrieval," *Appl. Soft Comput.*, vol. 125, Aug. 2022, Art. no. 109107.

[25] H. Zhang, W. Wu, K. Li, X. Zheng, X. Xu, X. Wei, and C. Zhao, "Multi-strategy active learning for power quality disturbance identification," *Appl. Soft Comput.*, vol. 154, Mar. 2024, Art. no. 111326.

[26] B. Quan, B. Yang, X. Li, Q. Chen, G. Liu, J. Hong, Z. Hao, L. Li, and H. Wei, "Research on the design of active learning algorithm based on query-by-committee for intelligent fetal monitoring," in *Proc. IEEE Int. Conf. Bioinf.*, May 2021, pp. 515–521.

[27] T. Evgeniou and M. Pontil, "Support vector machines: Theory and applications," in *Advanced Course on Artificial Intelligence*. Berlin, Germany: Springer, 1999, pp. 249–257.

[28] J. Piri and P. Mohapatra, "Exploring fetal health status using an association based classification approach," in *Proc. Int. Conf. Inf. Technol.*, 2019, pp. 166–171.

[29] A. K. Pradhan, J. K. Rout, A. B. Maharana, B. K. Balabantaray, and N. K. Ray, "A machine learning approach for the prediction of fetal health using ctg," in *Proc. 19th OITS Int. Conf. Inf. Technol.*, 2021, pp. 239–244.

[30] M. G. Signorini, N. Pini, A. Malovini, R. Bellazzi, and G. Magenes, "Integrating machine learning techniques and physiology based heart rate features for antepartum fetal monitoring," *Comput. Methods Programs Biomed.*, vol. 185, Mar. 2020, Art. no. 105015.

[31] E. Bahar, D. Agushinta R., Y. D. Arimbi, and M. Reksoprodjo, "Model structure of fetal health status prediction," *JUITA, Jurnal Informatika*, vol. 10, no. 2, p. 159, Nov. 2022.

[32] P. Warrick, E. Hamilton, R. E. Kearney, and D. Precup, "A machine learning approach to the detection of fetal hypoxia during labor and delivery," in *Proc. AAAI Conf. Artif. Intell.*, vol. 24, Jul. 2010, pp. 1865–1870.

[33] D. Campos and J. Bernardes, "Cardiotocography," UCI Mach. Learn. Repository, 2010. [Online]. Available: https://doi.org/10.24432/C51S4N

[34] J. Kaliappan, A. R. Bagepalli, S. Almal, R. Mishra, Y.-C. Hu, and K. Srinivasan, "Impact of cross-validation on machine learning models for early detection of intrauterine fetal demise," *Diagnostics*, vol. 13, no. 10, p. 1692, May 2023.

[35] S. C. R. Nandipati and X. Chew, "Classification and feature selection approaches for cardiotocography by machine learning techniques," *J. Telecommun., Electron. Comput. Eng. (JTEC)*, vol. 12, no. 1, pp. 7–14, Mar. 2020.

[36] P. Bhowmik, P. C. Bhowmik, U. A. M. Ehsan Ali, and M. Sohrawordi, "Cardiotocography data analysis to predict fetal health risks with tree-based ensemble learning," *Int. J. Inf. Technol. Comput. Sci.*, vol. 13, no. 5, pp. 30–40, Oct. 2021.

[37] Y. Fei, X. Huang, Q. Chen, J. Chen, L. Li, J. Hong, Z. Hao, and H. Wei, "Automatic classification of antepartum cardiotocography using fuzzy clustering and adaptive neuro-fuzzy inference system," in *Proc. IEEE Int. Conf. Bioinf. Biomed. (BIBM)*, Dec. 2020, pp. 1938–1942.

KAUSHAL BHARDWAJ (Member, IEEE) received the Ph.D. degree in computer science and engineering from Tezpur University, Tezpur, India, in 2020. He is currently an Assistant Professor with Bennett University, Greater Noida, Uttar Pradesh, India. His research interests include machine learning, natural language processing, mathematical morphology, and remote sensing image analysis. He serves as a Reviewer for various reputed journals, including IEEE Transactions on Geoscience and Remote Sensing, IEEE Transactions on Image Processing, and *International Journal of Remote Sensing*.

VANDNA SHARMA is currently pursuing the Bachelor of Technology (B.Tech.) degree in computer science engineering and technology with Bennett University, India. Her research interests include artificial intelligence, machine learning, and data science. This publication represents her initial foray into academic research and highlights her dedication to contributing to the field of computer science.

NIYATI GOYAL is currently pursuing the Bachelor of Technology (B.Tech.) degree in computer science engineering and technology with Bennett University, India. Her research interests include machine learning, data science, and computational intelligence. This publication marks her debut in academic research and it showcases her dedication to learning and contributing to the field of computer science.

BHAVIKA MITTAL is currently pursuing the Bachelor of Technology (B.Tech.) degree in computer science with Bennett University. Her research interests include artificial intelligence, machine learning, and data science. This publication signifies her first venture into academic research and underscores her commitment to advancing the field of computer science.

SHIV NARESH SHIVHARE (Senior Member, IEEE) received the Ph.D. degree in computer science and engineering from the Department of Computer Science and Engineering, National Institute of Technology at Uttarakhand, Srinagar, Uttarakhand, India. He is currently an Assistant Professor with the School of Computer Science Engineering and Technology, Bennett University (Times Group), Greater Noida, India. His research interests include computer vision and image processing, biomedical image analysis, machine intelligence, and classification.

○ ○ ○