Health and Technology (2023) 13:1–10
https://doi.org/10.1007/s12553-023-00732-8

REVIEW PAPER

Machine learning classification approach for asthma prediction models in children

Raphael Henshaw Ekpo[1] · **Victor Chukwudi Osamor**[1] · **Ambrose A. Azeta**[3] · **Excellent Ikeakanam**[1] · **Beatrice Opeyemi Amos**[2]

Received: 26 April 2022 / Accepted: 21 December 2022 / Published online: 31 January 2023

Abstract

Background Machine Learning refers to a methodology in the domain of data analytic that automates the systematic building of the model. It permits the discovery of unseen insights from an enormous datasets by means of suitable methods which involve repetitive learning gathered from data devoid of being programmed explicitly. The aim of this work is to explore machine learning strategies that are able to compensate with the weaknesses of existent asthma development predictive models in children. The objective of this review is to identify, compare, and summarize the existing machine and deep learning classification models for asthma prediction in children.
Methodology A substantial number of asthma development prediction models in children, such as conventional methods of risk factors, logical regression, and the hybrid of both statistical methods and risk factors existed. This study was performed following the guideline of Preferred Reporting Items for systematic Review and Meta Analysis (PRISMA). We carried a search for relevant studies from 2011-2021 using various online databases such as Google Scholar, Science Direct and PubMed on 23 July, 2021 to extract relevant papers on asthma prediction Models in children using machine learning and deep learning approaches.
Result The weaknesses associated with these existent asthma development predictive models in children include: they cannot be used as an appropriate tool for the implementation of decision support in electronic medical records, reduced clinical impact as well as low predictive accuracy. It was observed that ANN and SVM were among the best-performing algorithms in some machine learning comparative asthma prediction in children.
Conclusion This work concludes that there is a gradual increase of machine and deep learning algorithms for asthma prediction in children and that these approaches have shown greater predictive performance in pediatric asthma than the conventional existing models.

Keywords Machine learning · Deep learning · Performance metrics · Prediction accuracy · Pediatric asthma

✉ Victor Chukwudi Osamor
vcosamor@gmail.com;
victor.osamor@covenantuniversity.edu.ng

Raphael Henshaw Ekpo
ralphenshaw@gmail.com

Ambrose A. Azeta
aazeta@nust.na

Excellent Ikeakanam
excellent.ikeakanampgs@stu.cu.edu.ng

Beatrice Opeyemi Amos
opeamous@gmail.com

[1] Department of Computer and Information Sciences, Covenant University, Km 10, Idiroko Road, Canaanland, Ota, Ogun State, Nigeria

[2] Medical Laboratory Unit, State Hospital Ota, Onipanu, Ota, Ogun State, Nigeria

[3] Department of Software Engineering, Namibia University of Science and Technology(NUST), Windhoek, Namibia

1 Introduction

World Health Organization in 2020 regards asthma as one of the non-communicable chronic respiratory diseases that affect 339 million people globally with the higher number of deaths recorded in both low and middle lower economy nations. Four hundred and seventeen thousand, nine hundred and eighteen asthma-related deaths occurred in 2016. Asthma under diagnosis and under treatment is a significant problem not only to individual patients but also to their families and sometimes hinders individual activity throughout his lifespan [1–3]. Asthma is a common chronic disease marked by inflammation of the lung airways. It is heterogeneous in nature with distinguished symptoms like wheeze, shortness of breath, chest tightness, and coughs that change with time and strong effect, along with narrow exploratory airway flow [4].

A substantial number of asthma development prediction models in children, such as the conventional methods of risk factors, logical regression, and the hybrid of both statistical methods and risk factors existed. The weaknesses associated with these existent asthma development predictive models in children include: they cannot be used as an appropriate tool for the implementation of decision support in the electronic medical records, reduced clinical impact as well as low predictive accuracy, but the machine learning approaches has a greater advantage of addressing the attribute relatedness problem, making a distinction between important features and redundant features as well as increasing the prediction accuracy than the existing conventional methods [3, 5–9].

Machine Learning refers to a methodology in the domain of data analytic that automates the systematic building of the model. It permits the discovery of unseen insights from an enormous dataset by means of suitable methods which involve repetitive learning gathered from data devoid of being programmed explicitly. Machine Learning is an unusual and fast growing research domain which delivers effective and competent solutions by implementing suitable machine learning methods which changes from the old-fashioned approaches [10–12]. Machine learning has successfully been applied in several fields and in medical domain disease diagnosis and result prediction are the two major areas where machine learning techniques applications are of immense benefit [13, 14]. Machine learning has the capacity to take in several data to produce high prediction accuracy than the other traditional methods. They collect, analyze, and give different kinds of data in great ways, discovering trends and patterns. It also has the capacity to make a significant management of asthma such as exacerbation prediction, aids in reducing drug doses throughout the stable duration, tracing and reduction of cumulative corticosteroid doses as well as treatment of pathway triggering based on the condition of the weather [15].

Machine learning approaches overcome the restraints imposed by human mistakes and influence and are being used during the process of making decisions depending on the knowledge gathered from the input data. Moreover, machine learning techniques help to identify patients with great risk of asthma development prior to disease onset. This identification of great risk asthma patients will in turn make provision for timely prompt prevention and significantly mitigate asthma prevention [16]. Several traditional asthma prediction models in children have been investigated, but very limited studies are devoted to machine and deep learning classification approaches for asthma prediction models in children. This survey is necessary because of high global mortality rate of asthma according [1]. And asthma under-diagnosis and under-treatment are still clinical challenges in children [4]. Machine learning represents a novel approach to address the problems with the existing traditional asthma prediction model and will assist physicians in identifying asthmatic children for timely intervention as well as promote better healthcare delivery in pediatric asthma. The objective of this review is to identify, compare, and summarize the existing machine and deep learning classification models for asthma prediction in children. By identifying the various prediction models, we will attempt to answer the question to what extent are machine and deep learning models used for asthma prediction in children?. The remaining sections of this manuscript were organized as follows: Section 2 is the method used, Section 3 is the result and discussion and 4 is the conclusion.

2 Methods

2.1 Search strategy

This study was performed following the guideline of Preferred Reporting Items for systematic Review and Meta Analysis (PRISMA) [17]. We carried a search for relevant studies from 2011–2021 using various online databases such as Google Scholar, Science Direct and Pub Med on 23 July, 2021 to extract relevant papers on asthma prediction. Models in children using machine learning and deep learning approaches. The result of the search is shown in Fig. 1. The keywords used in this search process are as outlined as follows: "Machine Learning" OR " Deep Learning" OR "Artificial Intelligence" AND " Children with Asthma " OR "Pediatric Asthma" search strategy was used in PubMed and the same search strategy were equally performed in both Google Scholar as well as Science Direct.

2.2 Selection process

The inclusion selection criteria used for this work are enumerated as follows: the study must be asthma prediction, the study must use machine learning approach or deep learning, the study must be for asthma prediction in children, the study must state the evaluation metric used and must be done in the medical domain, the study must also be in a research article. The exclusion criteria for this work are those that do not use machine learning or deep learning, those that do not report performance metrics were also excluded, as well as papers that are written in a non-English language. The searched article titles, as well as the abstract, were screened manually and independently. Through discussion, disagreement on the articles to include was resolved by the authors. The entire text articles were reviewed and applied inclusion criteria to them. For the selected articles that meet the required criteria, the authors extracted information on the author, the purpose of prediction, feature selection, prediction model (methodology), population size, a total number of features associated with the dataset, no of the predictors used predictors, performance metrics, best performing algorithm., as well as patients age.

3 Result and discussion

From the search of the literature, 1373 articles were identified as shown in Fig. 1, 20 duplicate articles were removed. 1353 articles underwent abstract and title screening. The abstract and title screening exercise returned 157 articles,

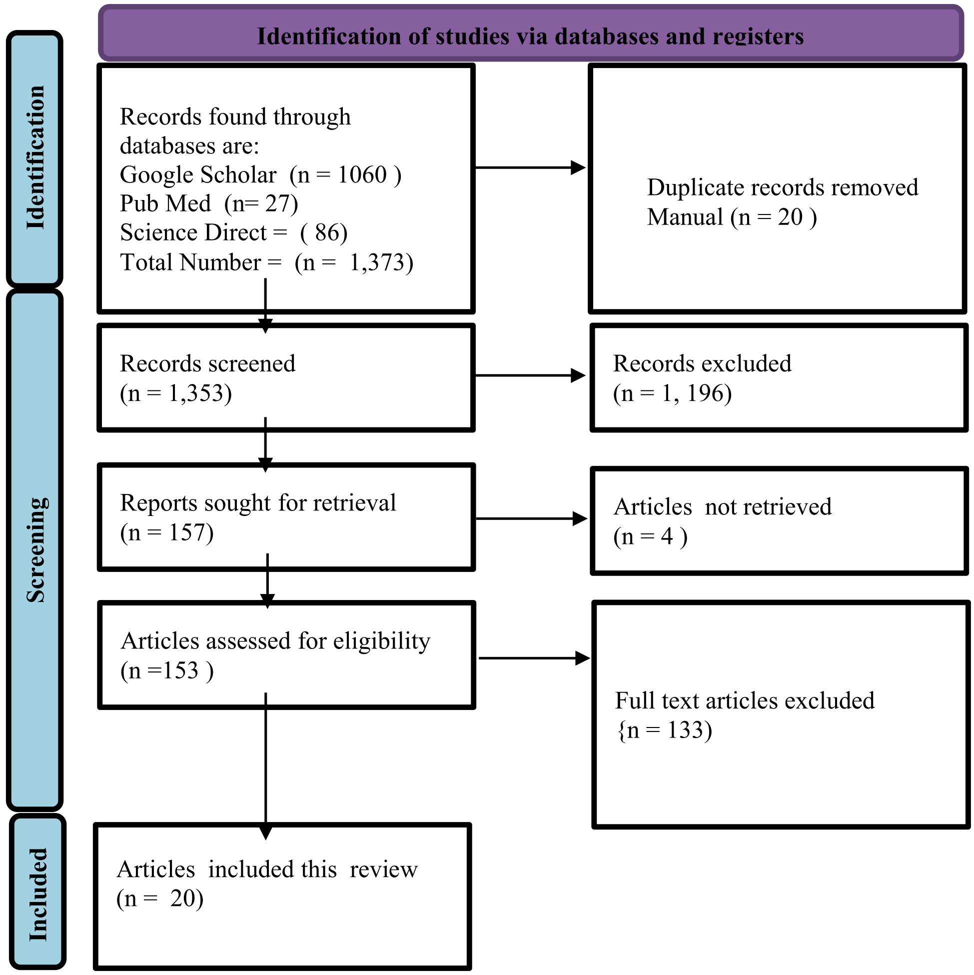

Fig. 1 PRISMA flow diagram for the study

of which 4 articles were not accessed, leaving 153 articles for full-text review. 20 articles were found to be relevant to this review. In this section of the review, the state-of-the-art machine and deep learning for asthma prediction in children are shown, as well as the related work in asthma.. The Summary of Machine learning Classification Approach Methodologies and Performance Metric Values for Asthma Prediction in children, described words and abbreviation used in the work as footer is shown in Table 1.

According to Prasadl et al. [18], the authors used clinical data and questionnaires to design an expert system for asthma diagnosis and comparison of various machine learning techniques. Auto associative memory, neural networks, Bayesian networks, Back propagation, C4.5 algorithm, and particle swarm optimization (PSO). For performance comparison and performance, study was done with sensitivity, specificity, and accuracy metrics. This study indicated that context-dependent associative memory model and PSO gave the best predicted outcome compared to the others. The authors used a limited number of questions that cannot capture all asthma diagnostics factors, thereby making the prediction accuracy questionable.

According to Chatzimichail et al.. [19], the authors focused on building predictive models for persistent asthma in children from seven years to fourteen. For feature extraction and classification of patterns, principal component analysis and support vector machine (SVM) classifier, respectively. Although the authors demonstrated the predictive ability of machine learning coupled with the benefits of using prognosis factors most of asthma already model, but it shows that there was a sign of over fitting in regard to the prediction result because of the limited size of the dataset used for the study. While cross-validation was for accuracy metrics for performance evaluation.

Chatzimichail et al. [20], study focused on the prediction of asthma persistent in diagnosed children age 5. The study feature extraction was carried out using partial least squares regression (PLSR) classifier, and classification was done using Artificial Neural Networks (ANN). The prediction metrics used are sensitivity, specificity, and classification accuracy. Although the authors demonstrated high predictive ability of machine learning coupled with the benefits of using large prognosis factors in the model, but it shows there was a sign of fitting in regard

Table 1 Summary of machine classification methodologies and performance metrics

Authors	Purpose of prediction	Feature selection method	Population size	Prediction model	No of factors associated with the model	NO of Most important features	ACC	Other Metrics			Best performing algorithm	AGE
Prasad et al. [18]	Expert system with machine learning for asthma disease diagnosis		25	CSAMM BN BP PSO C4.5			86.6 82.9 83.6 86.4 85.7				CSAMM & PSO	
Chatzimichail et al. [19]	Children Asthma Disease Persistent model	PCA	112	LSSVM	46	18	95.54					7–14
Chatzimichail et al. [20]	Children asthma Disease Persistent model	PLSR	112	MLP PNN	48	9	96.77 96.77	**SEN** 96.15 100	**SPE** 100 80		MLP and PNN	5–14 yrs
Princy et al. [21]	Comparison of several machine learning asthma prediction model			SVM MLPNN ADB MLP RF LSSVM PNN			98 80 90 98 90 96 98	**SEN** 96 80 90 96 90 100	**SPE** 96 80 90 100 90 80		SVM and MLP	
Spyroglou et al. [22]	Children asthma disease persistent prediction model	PCA	147	BLR	18	4–6	86.19	86.19	86.38			5-14yrs
Patel et al. [23]	Asthma hospitalization Prediction		29392	DT RF LLR GBM	8	8	72 82 83 84				GBM	2–18
Emanet et al. [24]	Machine learning Asthmatic patients Diagnosis prediction models		20	RF Adaboost ensemble ANN			90 80	**SEN** 90 80	SPE 90 80		RF and ADB Ensemble	
Harvey and Kumar [13]	Children asthma Development prediction	Correlation Matrix	50212	DT KNN RF NB LR	23	4	90.9				RF	0 -17 yrs
Ullah et al. [16]	Children asthma disease risk prediction		202	ANN RF SVM			92.1 92.1 94.1	**SEN** 94.7 97.3 98.0	SPE 84.6 90.9 82.7	**AUC** 84.6 90.9 82.7	SVM	

Table 1 (continued)

Authors	Purpose of prediction	Feature selection method	Population size	Prediction model	No of factors associated with the model	NO of Most important features	ACC	Other Metrics	Best performing algorithm	AGE
Xiaoyu et al. [25]	Asthma prediction Model for children	Statistical approach		BP ANN LR DT SVM		4	73.2	**SEN** 69.7 SPE 68	BP ANN	
Wang et al. [26]	Deep learning asthma disease risk prediction in Emergency Department		28378	LASSO LR ANN	33			**PRE** 0.123, 0.123 **F1 SCORE** 0.020, 0.198 **AUC** 84.2, 84.5 **RCC** 0.565, 0.510	ANN	6 months-18yrs
Kim et al. [27]	predicting the risk of asthma attack using deep learning		14	LTML MNL		7	PRE LTML more than 57–84% than MNL		LMTL	6–14
Chatzimichail et al. [29]	Children asthma disease persistent prediction model	GA	112	MPLNN	34	4		**AUC** 94.8		5–14
Ram et al. [30]	machine learning prediction model for the number of visits of asthma-related cases in emergency department	Backward Feature selection algorithm	4500 tweets from asthma disease stream	ANN DT ANN(Adaptive Boosting) Stacking(ANN + DT)		6	85.87	**PRE** 71.48 72.73 72.73 75.00	ANN	
Hogan et al. [31]	Asthma readmission prediction		1852	LR, CPH And ANN	5 9			**AUC** 0.52	ANN	5–18
Luo et al. [11]	developing prediction model for children asthma risk		210	MBDS SVM DL NB KNN RF			71.6 72.0 72.3 78.1 73.3 75.8	**SEN** 73.8, 71.5, 71.6, 58.8, 56.9, 48.3 **SPE** 71.4, 72.0, 72.5, 82.3, 77.0, 82.0 **RCC** 0.758, 0.718, 0.744, 0.777, 0.704, 0.662	MBDS	2–18
Lovrić et al. [32]	Children Asthma treatment prediction	Permutation importance	365	ADB RF	73	3	0.9698	**SEN** 0.9698 **SPE** 0.9661 **MCC** 0.9336	ADB	2–17, 18 -22

Table 1 (continued)

Authors	Purpose of prediction	Feature selection method	Population size	Prediction model	No of factors associated with the model	NO of Most important features	ACC	Other Metrics			Best performing algorithm	AGE
Bose et al. [33]	Asthma prediction in children	Anova -F, MultiSURF, Chi square, Relief as well as Embedded Feature selection	9934	NB KNN LR RF XGB	643	5	0.72 0.77 0.80 0.81 0.81	**F1 SCORE** 0.82 0.85 0.87 0.88 0.88	**PRE** 0.95 0.95 0.95 0.95 0.96	**RCC** 0.72 0.78 0.81 0.82 0.82	XGB	2–5, 5–10
Goto et al. [28]	Children Asthma prediction comparison	Lasso Regularization, Others inherent in modelling algorithm	2515	Hospitalization Traditional model reference, Lasso regression, GBDT RF DNN Critical care				**SEN** 83 67 74 74 71 54 71 78 76 78	**SPE** 55 75 71 71 74 91 86 77 84 83	**MCC** 55 75 71 71 74 91 86 77 84 83	GBDT and DNN	18 Yrs and below
Kothalawala et al. [34]	Children Asthma prediction	RFE with RF and univariant LR	1368	LSVM RDSVM DT NB RF MLP KNN	54	8 and 12 features for CAPE and CAPP models respectively		**F1 SCORE** 0.71 0.82	**AUC** 0.21 0.18		SVM with RBF in CAPE and SVM with Linear Kernel in CAPP model	10yrs

Where *RF* Random Forest, *ANN* Artificial Neural Network, *SVM* Support Vector Machine, *LR* Logistic Regression, *CPH* COX proportional Hazard, *DT* Decision tree, *GBM* Gradient Boost Machine, *LLR* Least Absolute Shrinkage and Selection Operator Logistic Regression, *PNN* Probability Neural Network, *LLP* LASSO Logistic Regression, *PR* Poison Regression, *LSTM* Long Short Term Memory, *DNN* Deep Neural Network, *MBDS* Multiboost with Decision Support, *LOAC* Level of Asthma Control, *FENO* Fractional Exhaled Nitric Oxide, *HSCRO* High sensitivity creative Protein, *CAPP* Children Asthma Prediction at Preschool, *SEN* Sensitivity, *AUC* Area Under the Receivers Operating Curve, *PLSR* Partial Least Square Regression, *CSAMM* Context sensitive Auto Association Memory Neural Network Model, *BN* Bayesian Network, *BP* Back propagation, *PSO* Particle Swarm Optimization, *LSSVM* Least Square Support Vector Machine, *ADB* Adaboost, *MLP* Multilayer Perceptron, *GBDT* Gradient Boosting Decision Tree, *NB* Naive Bayes, *MNL* Multinomial Logistic Regression, *GA* Genetic Algorithm, *RDSVM* Radial Bias kernel function Support Vector Machine, *MLPNN* Multilayer Perceptron Neural Networks, *FEVI* Forced Exploratory Volume in one seconds, *MEF50* Maximum Exploratory Flow of 50% of the vital flow capacity, *CAPE* Children Asthma Prediction in Early Life, *ACC* Accuracy, *PRE* Precision, *RCC* Recall, *MCC* Matthews Correlation Coefficient to the prediction result because of the limited size of the dataset used for the study.

According to Princy et al. [21], the authors used several machine learning classification algorithms for asthma prediction and compared the prediction accuracy with those models. Three breathing type test data was used for the study. The machine learning classification algorithms used by the authors are SVM, Multilayer Perceptron Neural Networks (MLPNN), Adaboost, Multilayer Perceptron (MLP), Random Forest (RF), LSSVM and Probabilistic Neural Network(PNN) and performance evaluation was done using the metrics of sensitivity, specificity as well as accuracy. The authors proposed that MLP and SVM be combined to form a single model for greater accuracy in future research. The gap in this study is that the authors did not indicate the category of patients used in the study.

According to Spyroglou et al. [22], the study focused on the evaluation of the accuracy of asthma persistent prediction in children using a mixture of Principal Component Analysis and Bayesian Logic Regression in clinical data. The authors proposed the model was able to efficiently and accurately predict asthma in a persistent in diagnosed asthmatic patients but was not directed to predicting asthma in children that have not been diagnosed with asthma.

Harvey et al. [13], developed asthma predictive models by analyzing National Survey of Children's health dataset using machine learning classifiers. The paper used Decision Tree, K-Nearest Neighbors (KNN), Random Forest, Linear Regression as well as Naive Bayes classifiers. The performance metrics used are sensitivity, specificity as well as accuracy. The result of the implementation indicated that Random Forest has the most prediction accuracy compared to other classifiers. The gap in this study dataset was not investigated for cases where class distribution examples are not equal.

Ullah et al. [16] applied a machine learning algorithm on asthmatic patient's dataset for prediction of asthma disease risk. The performance metrics used for this study are: sensitivity, specificity, accuracy, positive predictive value, Matthews Correlation Coefficient (MCC), negative predictive value, as well as receiver operator curve. Random Forest, SVM, and ANN were used as classification algorithms. The result of the study indicates that SVM outperforms the other classifiers based on the standard performance metrics used.

Patel et al. [23] focused on comparing machine learning models to predict the need for child asthma exacerbation for care in the hospital using the combination of triage clinical dataset in combination with neighborhood features as well as weather information of the patient. The performance of the study was carried out using area under the curve. The machine learning models used are Random forest, LASSO, regression, gradients, boost machine, and decision tree. The result indicated that the decision tree has the worst performance, while gradient boost machine slightly performs better than Random forest and LASSO regression logic.

Emanet et al. [24], developed predictive models using machine learning algorithms to diagnose asthmatic patients by using lung sound obtained from the patient's chest in a medical laboratory and comparing their predictive results. The authors used wavelength-based time–frequency data for prediction while Random Forest, Adaboost, and ANN were used for classification. The performance metric was carried out using the metric of specificity, sensitivity, and accuracy. The results of the study showed that the combination of Adaboost and random forest, and random forest algorithms give a better predictive accuracy of asthmatic patients than ANN. The limitation of these models is that the authors used a limited number of patients and only one prognostic factor which is the lung sound function, but asthma is a heterogeneous disease that comprises of many predictive factors for diagnosis and not a single factor even in adults.

Xiaoyu et al. [25], a study identifies variables that possess strong relationship with each other in childhood asthma from the database of Behavioral Risk Factor Surveillance System(BRFSS) and then build an asthma prediction model for children with back propagation artificial Neural Network (BPANN) using the required variables from the statistical approach on the dataset. This work limitation is the influence of the predictive accuracy may be encountered due to a great extent of missing information from the database used. The standard metric of accuracy, sensitivity, and specificity was used for the evaluation of the model.

Wang et al. [26] research work focused on the use of deep learning approaches to predict 3 months visits in the emergency department of asthma connected disease risk forecasting using the dataset of Medicaid claims. The result of the prediction was compared with the existent conventional statistical model of lasso logistic regression, and it shows that deep learning algorithm model of ANN which shows a better predictive accuracy than the Lasso Logistic Regression. The limitation of this work is that it could be strenuous to interpret the model and therefore require direct work with doctors for proper explanation, understanding, and interpretation of the model for the stakeholder's trust compared to other models.

Kim et al. [27] research focused on predicting the risk of asthma attack using deep learning methodology on performance exploratory flow data of 14 asthmatic children on a hospital visit. The results show that the long -short-term memory (LSTM) model was discovered to predict the classes of asthma risk attack better than Multinomial Logistic Regression (MNL). The limitation of the study is that the conclusions will be hard because of the limited sample size used for the study.

Goto et al. [28] focused on examining machine learning methodologies for clinical results of hospitalization and critical care with the traditional triage methodology reference model. The four machine learning models of Lasso, logistic regression (Lasso), gradient boosting decision tree (GBDT), random forest (RF), and deep neural networks (DNN). The results indicated that machine learning methodology models outperform the traditional triage model in hospitalization and critical care in specificity and sensitivity. Deep learning neural networks and gradient boosted decision tree give the highest net gain when compared to other machine learning models.

Chatzimichail et al. [29], the study focused on predicting asthma for children less than five years. The study feature selection was performed using a genetic algorithm and classification using Multilayer Perceptron Neural Networks (MLPNN). Classification accuracy was the performance metric used. Although the authors demonstrated the high predictive ability of the machine learning coupled model, it shows there was a sign of overfitting in regard to the prediction result because of the limited size of the dataset used for the study.

Ram et al. [30] focused on developing a machine learning prediction model for the number of visits of asthma-related cases in emergency departments using several data sources in a particular region. The machine learning models used are Artificial Neural Networks (ANN), Decision tree (DT), ANN (Adaptive boosting), as well as stacking (ANN + DT). The results indicated that the model can predict asthma-related emergency visit number with about 70% precision performance metric using social media and environmental data.

Hogan et al. [31] focused on comparing the conventional model with Artificial Neural Networks to identify major risk predictors for children's asthma readmission in six months. The result shows that the artificial neural network indicated nine risk factors with 0.637. Areas under the curve while five risk predictors were obtained from logistic regression and Cox proportional hazard with 0.592 area under the operation curve. Hogan et al. [31], concluded that we can have different asthma readmission models, but relying on the conventional model alone will lead to overlooking major useful risk predictors which can be identified by other machine learning like artificial neural networks.

Luo et al. [11] focused on developing a prediction model for children's asthma risk control in a week time before it occurs, using clinical information, demographic information as well as environmental information combined. Multi boost Decision stump model has the best result using the standard performance metric of accuracy, sensitivity, area under the receiver characteristics curve.

According to Lovrić et al. [32], the work focuses on the prediction of children's asthma treatment outcomes as well as identifying the significant features used to understand the basic process by using a machine learning model. The outcome of using machine learning shows that asthma control, as well as Fractional Exhaled Nitric Oxide (FENO) lung function-based, has a better prediction accuracy than both Forced Exploratory Volume in One Second (FEV1) and Maximum Exploratory Flow of 50% of the vital flow capacity(MEF50) in children, asthma treatment outcome prediction, asthma severity, total, *Immunoglobulin E* (IgE) and High sensitivity Creative Protein (hsCRP) are the most important predictive variables for children asthma patient treatment prediction outcome.

Bose et al. [33] focused on developing machine learning prediction models to predict transient asthma and persistent asthma in children. The models used Naive Bayes, Random Forest, Logistic Regression, K Nearest Neighbour as well as Gradient Boost Machine algorithms. The outcome indicated that gradient boosting machines outperform other algorithms. The model's most important variables are age of the last under 5, asthma diagnosis, the total number of related asthma visits, self-identified black race, allergic rhinitis, as well as eczema. According to Kothalawala et al. [34], the study focused on using machine learning techniques to predict asthma disease in both preschool and school ages in CAPE and CAPP models. The outcome indicates that SVM with RBF outperformed other machine learning algorithms in the CAPE model with 0.71 and 0.21 prediction performances in AUC and Brier score, respectively. On the other hand, SVM with linear kernel also outperformed other machine learning algorithms in the CAPP model with 0.82 and 0.18 prediction performances in both AUC and Brier scores respectively. The two machine learning models also demonstrated better generality than the conventional regression model for asthma prediction in children.

A total of 32 machine learning algorithms have been used for asthma prediction in children. 8 studies used SVM and its variants algorithm, 7 works used RF, 4 studies used DT and ANN, 3 studies used NB, LR, KNN, LLR as well as MLP, 2 works used XGB, DNN, MLPNN, and ADB. Seven studies used large population sizes ranging from 1368–50212, while eight (8) studies used small population sizes between 14–365.

Ten (10) works used different feature selection methods and eight (8) studies out of the 20 used a large number of factors associated with a model prediction while 3 of the studies used a small number of factors with the model prediction. 7 studies used 3–6 most important variables after feature selection for the model developments, while three (3) studies used between 7- 18 most important features for the model development. Only one study [34] externally validated the work of another asthma study.

Four (4) out of twenty (20) studies compare deep neural networks with other machine learning algorithms and also with traditional regression method, DNN outperformed in 3 studies with the exception of one study [27].. 3 works in this review included algorithm explanability /interpretability. These works are [6, 32, 35], Luo [6], developed an automatic explanation approach to differentiate algorithm explanation from algorithm prediction. It performs these actions by employing two models consecutively; the first one for prediction and rule-based associative classifier method for algorithm explanation. Lovrić et al. [32] employed permutation importance to understand the most significant features for separating responders as well as non responders to asthma treatment success. Alsaad et al. [35] employed conceptual decomposition interpretability approach to generate relevant score from both LSTM and bidirectional LSTM, as well as to predict future medical result utilizing Electronic Heath Record dataset.

The study indicated that there is a gradual increase of machine learning and deep learning algorithm for asthma prediction in children, although much research is still required in this domain. It was also observed that ANN and SVM were among the best-performing algorithms in some machine learning comparative asthma prediction in children; this collaborates with a similar trend as kidney stone prediction using machine learning algorithms shown in [12]. Other best performing algorithms are GBM and DNN despite their limited use in asthma prediction models in children This study indicated that machine learning and deep learning algorithms consistently outperform the traditional logistic regression in asthma prediction in children in performance generally. Although a couple of complex interpretability concerns such as trust, moral and fairness issues are associated with these black box algorithms in medical domain, there is need for more research in interpretability models for these algorithms in asthma prediction in children. This will form the basis for further study.

4 Conclusion

Machine learning has successfully been applied in several fields, and in the medical domain, disease diagnosis and result prediction are the two significant areas where machine learning techniques are of immense benefit. The authors of this work specifically explore the machine learning classification approach for asthma prediction in children. The study indicated that there is a gradual increase in machine and deep learning algorithms for asthma prediction in children. In conclusion, machine learning and deep learning have shown greater predictive performance in pediatric asthma than the conventional existing asthma models; however, more research is still expected to be done in the machine learning and deep learning asthma prediction domain in children, especially in the use of deep learning algorithms.

Authors' contributions VCO conceived the idea of the work. VCO, AAA, RHE, BOA and EI designed, analyzed and wrote the manuscript. Revision and editing of the manuscript was done by all the authors while VCO and AAA supervised the entire work.

Funding The authors declare that this work was done through self-funding.

Data availability Not applicable.

Code availability Not applicable.

Declarations

Ethics approval Not applicable.

Consent to participate Not applicable.

Consent for publication Not applicable).

Conflict of interest/competing interests On behalf of all authors, the corresponding author states that there is no conflict of interest.

References

1. World Health Organization asthma. Fact sheet. retrieved from http://www.who.int/respiratory/asthma/en/. Accessed Feb 2020.
2. Kikidis D, Votis K, Tzovaras D. MyAirCoach: Designing a mobile application for the education of patients regarding asthma disease. In: 2015 International Conference on Interactive Mobile Communication Technologies and Learning (IMCL). IEEE; 2015. p. 373–377.
3. Kothalawala DM, Kadalayil L, Weiss VB, et al. Prediction models for childhood asthma: a systematic review. Pediatr Allergy Immunol. 2020;31(6):616–27.
4. Reddel HK, Bateman ED, Becker A, et al. A summary of the new GINA strategy: a roadmap to asthma control. Eur Respir J. 2015;46(3):622–39.
5. Luo G, Nkoy FL, Stone BL, Schmick D, Johnson MD. A systematic review of predictive models for asthma development in children. BMC Med Inform Decis Mak. 2015;15(1):1–16.
6. Luo G. Automatically explaining machine learning prediction results: a demonstration on type 2 diabetes risk prediction. Health Inf Sci Syst. 2016;4(1):1–9.
7. Luo G, Sward K. A roadmap for optimizing asthma care management via computational approaches. JMIR Med Inform. 2017;5(3):e32.
8. Cai J, Luo J, Wang S, Yang S. Feature selection in machine learning: A new perspective. Neurocomputing. 2018;300:70–9.
9. Jain D, Singh V. Feature selection and classification systems for chronic disease prediction: A review. Egypt Inform J. 2018;19(3):179–89.
10. Desarkar A, Ajanta D. Big-Data analytics, machine learning algorithms and scalable /parallel /distributed algorithms. In: Internet of things and big data technologies for next generation healthcare. Springer, Cham 2017; 23:159–197.
11. Luo G, Stone BL, Fassl B, Maloney CG, Gesteland PH, Yerram SR, Nkoy FL. Predicting asthma control deterioration in children. BMC Med Inform Decis Mak. 2015;15(1):1–8.
12. Jayatilake SMDAC, Ganegoda GU. Involvement of machine learning tools in healthcare decision making. J Healthc Engi. 2021;1–20.
13. Harvey JL, Kumar SA. Machine learning for predicting development of asthma in children. In: IEEE symposium series on computational intelligence; 2019. p. 596–603.
14. Sidey-Gibbons JA, Sidey-Gibbons CJ. Machine learning in medicine: a practical introduction. BMC Med Res Methodol. 2019;19(1):1–18.

15. Messinger AI, Luo G, Deterding RR. The doctor will see you now: How machine learning and artificial intelligence can extend our understanding and treatment of asthma. J Allergy Clin Immunol. 2020;145(2):476–8.
16. Ullah R, Khan S, Ali H, et al. A comparative study of machine learning classifiers for risk prediction of asthma disease. Photodiagn Photodyn Ther. 2019;28(1):292–6.
17. Page MJ, McKenzie JE, Bossuyt PM, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ. 2021;2021:372.
18. Prasadl BD, Prasad PE, Sagar Y. An approach to develop expert systems in medical diagnosis using machine learning algorithms (asthma) and a performance study. Int J Soft Comput. 2011;2(1):26–33.
19. Chatzimichail E, Paraskakis E, Sitzimi M, Rigas A. An intelligent system approach for asthma prediction in symptomatic preschool children. Comput Math Methods Med. 2013:1–5.
20. Chatzimichail E, Paraskakis E, Rigas A. Predicting asthma outcome using partial least square regression and artificial neural networks. Adv Artif Intell. 2013;16877470:1-7.
21. Princy JC, Sivaranjani K. Survey on asthma prediction using classification technique. Int J Comput Sci Mob Comput. 2016;5(1):515–8.
22. Spyroglou II, Spöck G, Chatzimichail EA, Rigas A, Paraskakis E. A Bayesian logistic regression approach in asthma persistence prediction. Epidemiol Biostat Public Health. 2018;15(1):e12777.
23. Patel SJ, Chamberlain DB, Chamberlain JM. A machine learning approach to predicting need for hospitalization for pediatric asthma exacerbation at the time of emergency department triage. Acad Emerg Med. 2018;25(12):1463–70.
24. Emanet N, Öz HR, Bayram N, Delen D. A comparative analysis of machine learning methods for classification type decision problems in healthcare. Decis Anal. 2014;1(1):1–20.
25. Xiaoyu M, Zhang H, Zhao Y. Building childhood asthma prediction model with artificial neural network and BRFSS database. Data Anal Knowl Disc. 2018;2(8):10–5.
26. Wang X, Wang Z, Pengetnze YM, Lachman BS, Chowdhry V. Deep learning models to predict pediatric asthma emergency department visits. arXiv preprint arXiv:1907.11195. 2019;1–7.
27. Kim D, Cho S, Tamil L, Song DJ, Seo S. Predicting asthma attacks: effects of indoor PM concentrations on peak expiratory flow rates of asthmatic children. IEEE Access. 2019;8:8791–7.
28. Goto T, Camargo CA, Faridi MK, Freishtat RJ, Hasegawa K. Machine learning–based prediction of clinical outcomes for children during emergency department triage. JAMA Netw Open. 2019;2(1):e186937.
29. Chatzimichail E, Paraskakis E, Rigas A. An Evolutionary Two-Objective Genetic Algorithm for Asthma Prediction UKSim 15th International Conference on Computer Modelling and Simulation: IEEE; 2013. p. 90–94.
30. Ram S, Zhang W, Williams M, Pengetnze Y. Predicting asthma-related emergency department visits using big data. IEEE J Biomed Health Inform. 2015;19(4):1216–23.
31. Hogan AH, Brimacombe M, Mosha M, Flores G. Comparing Artificial intelligence and traditional methods to identify factors associated with pediatric asthma readmission. Acad Pediatr. 2021;22(1):55–61.
32. Lovrić M, Banić I, Lacić E, Pavlović K, Kern R, Turkalj M. Predicting treatment outcomes using explainable machine learning in children with asthma. Children. 2021;8(5):376.
33. Bose S, Kenyon CC, Masino AJ. Personalized prediction of early childhood asthma persistence: A machine learning approach. PLoS ONE. 2021;16(3):e0247784.
34. Kothalawala DM, Murray C, Simpson A, Custovic A, et al. Development of childhood asthma prediction models using machine learning approaches. Clin Transl Allergy. 11(9):e12076.
35. AlSaad R, Malluhi Q, Janahi I, Boughorbel S. Interpreting patient-Specific risk prediction using contextual decomposition of BiLSTMs: application to children with asthma. BMC Med Inform Decis Mak. 2019;19(1):1–1.