IJCSE International Journal of Computer Sciences and Engineering Open Access
Survey Paper Vol.-7, Issue-5, May 2019 E-ISSN: 2347-2693

Applications of Machine Learning Techniques on Prediction of Children's various health problems: A Survey

A. Meharaj Begum

Dept. of CS, AIMAN College of Arts and Science for Women,Bharathidasan University, Tiruchirappalli, India.

DOI: https://doi.org/10.26438/ijcse/v7i5.11661176 | Available online at: www.ijcseonline.org

Accepted: 19/May/2019, Published: 31/May/2019

***Abstract* -** Early diagnosis of children's health problems helps the professionals to treat it at an earlier stage and improves their quality of life. The life skills of the young minds are the best investment to build the healthy society of the Country. Researchers across various countries are interested to predict health problems related to children with respect to parenting style, hereditary health issues, food habits and physical activities. Machine learning plays a vital role in analyzing and predicting the hidden facts in the data we collected. The main objective of this study is to present an overview of many machine learning techniques such as Support Vector Machines, Naive Bayes classifier, K-Nearest Neighbor, Decision Tree, K-means algorithm and perform a comparative analysis of their accuracy and help the researchers to choose best algorithm on prediction of Children's various health problems such as Early Childhood Obesity, Anxiety Disorders, Attention Deficit Hyperactive Disorder, Mental Health Problems, Child Post Traumatic Stress, Autism Spectrum Disorder and Insulin Resistance in Children. This survey paper can lead to develop innovative and efficient algorithms on prediction of children's health problems to improve their quality of life in a better way.

Keywords: Machine Learning Algorithms, Prediction, Children's Health Problems.

I. INTRODUCTION

Machine learning plays an extensive role for development especially in terms of data analytics. Health care is one of the fields where Machine Learning made big phases of enhancements because of the huge amount of data being processed and analyzed.

The three categories of these Machine Learning algorithms are:

- Supervised Learning
- Unsupervised Learning
- Reinforcement Learning

A. Supervised Learning

Supervised learning is the machine learning task of learning a function that maps an input to an output based on example input-output pairs. It infers a function from labeled training data consisting of a set of training examples. In supervised learning, each example is a pair consisting of an input object and a desired output value. And while using Training dataset, the process can be thought of as a teacher Supervising the Learning Process. The correct answer is known and stored in the system already. The algorithm helps in making Predictions about the Data that is in Training Process and gets the correction done by the Teacher itself. There is an end to the learning only when the Algorithm has achieved an acceptable degree or level of Performance.

There are two types of Supervised learning problems:

- **Classification Problems:** Classification problem can be defined as the problem that brings output variable which falls just in particular categories, such as the "red" or "blue" or it could be "disease" and "no disease".
- **Regression**: A regression problem is when the output variable is a real value, such as "dollars" or it could be "weight".

There are few supervised machine learning algorithms, such as:

- Support vector machines for classification problems
- K-Nearest Neighbor
- Naive Bayes Classification
- Decision Trees
- Random forest for classification and regression problems
- Linear regression
- Ordinary Least Squares Regression
- Logistic Regression
- Ensemble Methods

1). Support vector machine

The support vector clustering algorithm was created to categorize unlabeled data, and is one of the most widely used clustering algorithms in industrial applications. A support vector machine constructs a hyperplane or set of hyperplanes in a infinite-dimensional space, which can be used for classification, regression, or the tasks like outliers detection. Intuitively, a good separation is achieved by the hyperplane that has the largest distance to the nearest training-data point of any class (so-called functional margin), since in general the larger the margin the lower the generalization error of the classifier.

Applications of SVM

- SVMs are helpful in text and hypertext categorization as their application can significantly reduce the need for labeled training instances in both the standard inductive and transductive settings.
- Classification of images can also be performed using SVMs. Experimental results show that SVMs achieve significantly higher search accuracy than traditional query refinement schemes after just three to four rounds of relevance feedback. This is also true of image segmentation systems, including those using a modified version SVM that uses the privileged approach as suggested by Vapnik.
- Hand-written characters can be recognized using SVM.
- The SVM algorithm has been widely applied in the biological and other sciences. They have been used to classify proteins with up to 90% of the compounds classified correctly. Permutation tests based on SVM weights have been suggested as a mechanism for interpretation of SVM models. Support vector machine weights have also been used to interpret SVM models in the past. Post hoc interpretation of support vector machine models in order to identify features used by the model to make predictions is a relatively new area of research with special significance in the biological sciences.

2). K-Nearest Neighbor

In pattern recognition, the *k*-nearest neighbors algorithm (*k*-NN) is a non-parametric method used for classification and regression. In both cases, the input consists of the *k* closest training examples in the feature space. The output depends on whether *k*-NN is used for classification or regression:

- In k-NN classification, the output is a class membership. An object is classified by a majority vote of its neighbors, with the object being assigned to the class most common among its *k* nearest neighbors (*k* is a positive integer, typically small). If $k = 1$, then the object is simply assigned to the class of that single nearest neighbor.
- In k-NN regression, the output is the property value for the object. This value is the average of the values of its *k* nearest neighbors.
- *k*-NN is a type of instance-based learning, or lazy learning, where the function is only approximated locally and all computation is deferred until classification. The *k*-NN algorithm is among the simplest of all machine learning algorithms.
- Both for classification and regression, a useful technique can be used to assign weight to the contributions of the neighbors, so that the nearer neighbors contribute more to the average than the more distant ones. For example, a common weighting scheme consists in giving each neighbor a weight of $1/d$, where d is the distance to the neighbor.
- The neighbors are taken from a set of objects for which the class (for *k*-NN classification) or the object property value (for *k*-NN regression) is known. This can be thought of as the training set for the algorithm, though no explicit training step is required.
- A peculiarity of the *k*-NN algorithm is that it is sensitive to the local structure of the data. The algorithm is not to be
- confused with *k*-means, another popular machine learning technique.

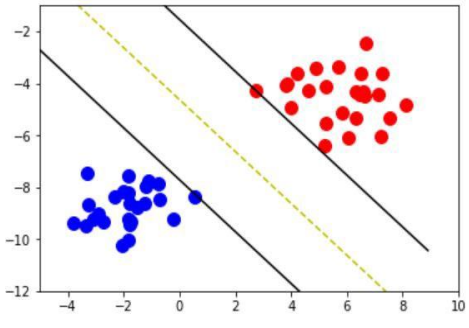


Figure 1 Hyper plane of SVM

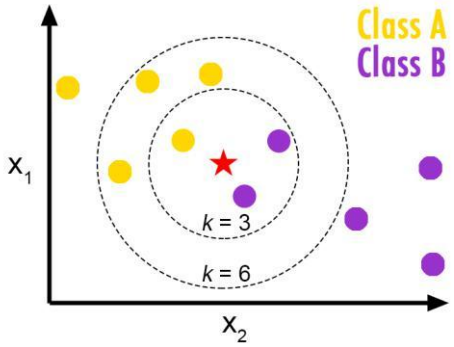


Figure 2 k-NN classification

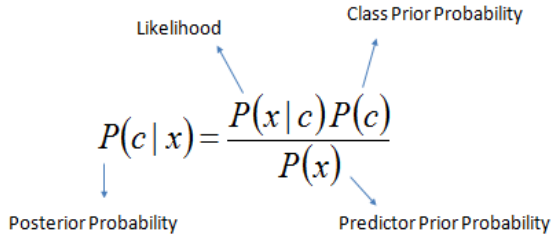
3). Naive Bayesian

In machine learning, naive Bayes classifier are a family of simple "probabilistic classifiers" based on applying Bayes' theorem with strong(naive) independence assumptions between the features.

Naive Bayes has been studied extensively since the 1950s. It was introduced under a different name into the text retrieval community in the early 1960s, and remains a popular (baseline) method for text categorization, the problem of judging documents as belonging to one category or the other (such as spam or legitimate, sports or politics, etc.) with word frequencies as the features. With appropriate pre-processing, it is competitive in this domain with more advanced methods including support vector machines. It also finds application in automatic medical diagnosis.

Naive Bayes classifiers are highly scalable, requiring a number of parameters linear in the number of variables (features/predictors) in a learning problem. Maximum-likelihood training can be done by evaluating a closed-form expression, which takes linear time, rather than by expensive iterative approximation as used for many other types of classifiers.

In the statistics and computer science literature, naive Bayes models are known under a variety of names, including simple Bayes and independence Bayes. All these names reference the use of Bayes' theorem in the classifier's decision rule, but naive Bayes is not a Bayesian method.



$$P(c \mid \mathrm{X}) = P(x_1 \mid c) \times P(x_2 \mid c) \times \cdots \times P(x_n \mid c) \times P(c)$$

Figure 3 Naïve Bayesian Probability

Bayes theorem provides a way of calculating the posterior probability, $P(c|x)$, from $P(c)$, $P(x)$, and $P(x|c)$. Naive Bayes classifier assume that the effect of the value of a predictor (x) on a given class (c) is independent of the values of other predictors. This assumption is called class conditional independence.

4). Decision Tree

Decision tree is the most powerful and popular tool for classification and prediction. A Decision tree is a flowchart like tree structure, where each internal node denotes a test on an attribute, each branch represents an outcome of the test, and each leaf node (terminal node) holds a class label.

Construction of Decision Tree:

A tree can be *"learned"* by splitting the source set into subsets based on an attribute value test. This process is repeated on each derived subset in a recursive manner called *recursive partitioning*. The recursion is completed when the subset at a node all has the same value of the target variable, or when splitting no longer adds value to the predictions. The construction of decision tree classifier does not require any domain knowledge or parameter setting, and therefore is appropriate for exploratory knowledge discovery. Decision trees can handle high dimensional data. In general decision tree classifier has good accuracy. Decision tree induction is a typical inductive approach to learn knowledge on classification.

Decision Tree Representation:

Decision trees classify instances by sorting them down the tree from the root to some leaf node, which provides the classification of the instance. An instance is classified by starting at the root node of the tree, testing the attribute specified by this node, then moving down the tree branch corresponding to the value of the attribute as shown in the Figure 4. This process is then repeated for the subtree rooted at the new node.

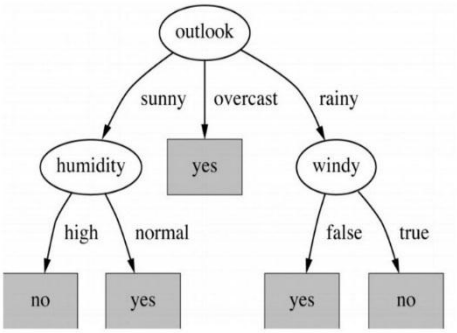


Figure 4 Decision Tree

The decision tree shown in Figure 4 classifies a particular morning according to whether it is suitable for playing tennis and returning the classification associated with the particular leaf.(in this case Yes or No). For example, the instance

(Outlook = Rain, Temperature = Hot, Humidity = High, Wind = Strong)

would be sorted down the leftmost branch of this decision tree and would therefore be classified as a negative instance.

In other words, we can say that decision tree represents a disjunction of conjunctions of constraints on the attribute values of instances.

(Outlook = Sunny ^ Humidity = Normal) v (Outlook = Overcast) v (Outlook = Rain ^ Wind = Weak)

Strengths and Weakness of Decision Tree approach
The strengths of decision tree methods are:

- Decision trees are able to generate understandable rules.
- Decision trees perform classification without requiring much computation.
- Decision trees are able to handle both continuous and categorical variables.
- Decision trees provide a clear indication of which fields are most important for prediction or classification.

The weaknesses of decision tree methods:

- Decision trees are less appropriate for estimation tasks where the goal is to predict the value of a continuous attribute.
- Decision trees are prone to errors in classification problems with many class and relatively small number of training examples.
- Decision tree can be computationally expensive to train. The process of growing a decision tree is computationally expensive. At each node, each candidate splitting field must be sorted before its best split can be found. In some algorithms, combinations of fields are used and a search must be made for optimal combining weights. Pruning algorithms can also be expensive since many candidate sub-trees must be formed and compared.

B. Unsupervised Learning

Unsupervised learning is that algorithm where we only have to insert/put the input data (X) and no corresponding output variables are to be put. The major goal for the unsupervised learning is to help model the underlying structure or maybe in the distribution of the data in order to help the learners learn more about the data. These are termed as unsupervised learning because unlike supervised learning there are no correct answers and there is no teacher to this. Algorithms are left to their own devices to help discover and present the interesting structure that is present in the data.

Unsupervised learning problems can even be grouped ahead into clustering and association problems.

- **Clustering:** A clustering is that problem which indicates what you want to discover and this helps in the inherent groupings of the data, such as grouping the customers based on their purchasing behavior.
- **Association:** An association rule is termed to be the learning problem. This is where you would be discovering the exact rules that will describe the large portions of your data. Example: People who buy X are also the one who tends to buy Y.

Some examples of unsupervised learning algorithms are:

- K-means for clustering problems
- Apriori algorithm for association rule learning problems
- Principal Component Analysis
- Singular Value Decomposition
- Independent Component Analysis

K-means clustering

K-means clustering is a method used for clustering analysis, especially in data mining and statistics. It aims to partition a set of observations into a number of clusters (k), resulting in the partitioning of the data into Voronoi cells. It can be considered a method of finding out which group a certain object really belongs to.

It is used mainly in statistics and can be applied to almost any branch of study. For example, in marketing, it can be used to group different demographics of people into simple groups that make it easier for marketers to target. Astronomers use it to sift through huge amounts of astronomical data; since they cannot analyze each object one by one, they need a way to statistically find points of interest for observation and investigation.

The algorithm:

- K points are placed into the object data space representing the initial group of centroids.
- Each object or data point is assigned into the closest k.
- After all objects are assigned, the positions of the k centroids are recalculated.
- Steps 2 and 3 are repeated until the positions of the centroids no longer move

The main contribution of this study is to provide an idea to medical analysts and researchers to choose better algorithms which gives better accuracy and reliability to predict specific health problems. It also provides basis for the improvement of future researches on prediction of other health problems.

II. RELATED WORK

A. Prediction of Early Childhood 'Obesity'

Dugan, T. M., et al (2015), aims to predict childhood obesity after age two, using only data collected prior to the second birthday by a clinical decision support system called CHICA. Analysis of six different machine learning methods: **Random Tree, Random Forest, J48, ID3, Naïve Bayes, and Bayes** trained on CHICA data show that an accurate, sensitive model can be created. Of the methods analyzed, the ID3 model trained on the CHICA dataset proved the best overall performance with accuracy of 85% and sensitivity of 89%. Additionally, the ID3 model had a positive predictive value of 84% and a negative predictive value of 88%. The structure of the tree also gives insight into the strongest predictors of future obesity in children. Many of the strongest predictors seen in the ID3 modeling of the CHICA dataset have been independently validated in the literature as correlated with obesity, thereby supporting the validity of the model [1].

B. Prediction of Children's 'Anxiety Disorders'(GAD)

Theodore, K., et al., (2018) defines that Anxiety disorders are a type of mental disorders characterized by important feelings of fear and anxiety. In this work, a comparative literature search was conducted on research for the prediction of specific types of anxiety disorders using machine learning techniques. Several machine learning techniques have been utilized to develop accurate prediction models in order to assist in providing better medical services. But Hybrid methods and SVM were the most highly used method especially for the prediction of GAD. Moreover, ANNs and ensemble methods performed very well, managing to achieve the highest prediction scores while NFS had some good scores too. Furthermore, we saw that the combination of multi-level features using multi-kernel learning can lead to an improvement of the classification performance for predicting GAD [2].

C. Predicting 'Attention Deficit Hyperactive Disorder' (ADHD)

Shah, A. (2015) proposes a system to predict whether the patient will have ADHD or not. It is a challenging task. Diagnosing ADHD is difficult during the initial stages because many of the early symptoms can be caused by other more common health conditions. There's no single test for ADHD, and diagnosis is done by observing person's symptoms and behavior. There can be delays in diagnosis as sometimes the doctors themselves need some time and observation to be sure of the situation, particularly in cases where the condition progresses very slowly. Thus, there is a need for mechanism that can accurately detect ADHD. It is a Machine Learning problem, chosen to be solved using Supervised Learning techniques. The problem of detection of whether or not a child has ADHD belongs to binary classification domain where the result is either yes or no. Dataset used in the project is obtained from Vaccine Adverse Event Reporting System (VAERS) which is a nation vaccine surveillance program which collects and shares information about adverse events following immunization. Dataset contains patient's symptom stories. The problem is the presence of random 10 textual data and phrases, which makes the task of prediction very challenging. Moreover, the dataset requires being cleaned, filtered and restructured to make it useful. Special characters, language specifics and repetitions in data need to be dealt with carefully. Naive Bayes and Logistic Regression models have been built in Weka and Apache Spark. Model based on logistic regression performs better than model built using Naive Bayes classifier. Hence, if the model is used in real world scenario then it can prove to solve some diagnostic problems of ADHD [3].

D. Prediction of Mental Health Problems Amon Children Using Machine Learning Techniques

Sumathi, M. R., & Poorna, B. (2016) stipulates that nowadays, a number of expert systems are utilized in medical domain to predict diseases accurately at an early stage so that treatment can be made effectively and efficiently. Also, expert systems are developed in the mental health domain to predict the mental health problem at an earlier stage. As a number of machine learning techniques are available to construct expert systems, it is necessary to compare them and identify the best that suits the domain of interest. The research has compared eight machine learning techniques (classifiers) on classifying the dataset to different mental health problems. It is evident from the results that the three classifiers viz., Multilayer Perceptron, Multiclass Classifier and LAD Tree produce more accurate results than the others. The data set is very minimal and in future, the research may be applied for a large data set to obtain more accuracy. The classifiers need to be trained prior to the implementation of any technique in real prediction [4].

E. Prediction of short-term outcome in 'Autism Spectrum Disorder' (ASD)

Usta, M. B., et al (2018), proposes a system to examine the predictors of outcome with machine learning methods, which are novel computational methods including statistical estimation, information theories and mathematical

learning automatically discovering useful patterns in large amounts of data. They study the group comprised 433 children (mean age: 72.3 ± 45.9 months) with ASD diagnosis. The ASD symptoms were assessed by the Autism Behavior Checklist, Aberrant Behavior Checklist, Clinical Global Impression scales at baseline (T0) and 12th (T1), 24th (T2), and 36th (T3) months. They tested the performance of for machine learning algorithms (Naive Bayes, Generalized Linear Model, Logistic Regression, Decision Tree) on our data, including the 254 items in the baseline forms. Patients with ≤2 CGI points in ASD symptoms at in 36 months were accepted as the group who has "better outcome" as the prediction class. The significant proportion of the cases showed significant improvement in ASD symptoms (39.7% in T1, 60.7% in T2; 77.8% in T3). Our machine learning model in T3 showed that diagnosis group affected the prognosis. In the autism group, older father and mother age; in PDD-NOS group, MR comorbidity, less birth weight and older age at diagnosis have a worse outcome. In Asperger's Disorder age at diagnosis, age at first evaluation and developmental cornerstones has affected prognosis [5].

F. Using Multiple Machine Learning Algorithms to Predict Autism in Children

Alarifi, H. S., & Young, G. S. (2018) aims to implement and compare machine learning techniques to develop a model that can predict Autism Spectrum Disorder (ASD). Autism Spectrum Disorder is a developmental and neurological disorder. Autism is also identified as a range of conditions categorized by various challenges such as social skills, repetitive behaviors, speech and non-verbal communication. Autistic children and adults have their unique strengths and differences. The goal is to measure accuracy for multiple methods to better assess, and then develop a model that is used predict Autism among children. This is by applying traditional Autism test for children AQ-10, which is commonly used by psychologists and pediatricians to diagnose autism. Based on the information collected from AQ-10 tests for 292 children. This work is able to compare several Linear and Non-linear classifiers and neural network. Finally, it is clear that logistic regression did the best performance [6].

G. Machine learning methods to predict child Post Traumatic Stress: a proof of concept study

Saxe, G. N., et al., (2017) discuss with Machine Learning predictive classification methods – with causal discovery feature selection – were applied to a data set of 163 children hospitalized with an injury and PTSD was determined three months after hospital discharge. At the time of hospitalization, 105 risk factor variables were collected spanning a range of biopsychosocial domains. Seven percent of subjects had a high level of PTSD symptoms. A predictive classification model was discovered with significant predictive accuracy. A predictive model constructed based on subsets of potentially causally relevant features achieves similar predictivity compared to the best predictive model constructed with all variables. Causal Discovery feature selection methods identified 58 variables of which 10 were identified as most stable. In this first proof-of-concept application of ML methods to predict childhood Posttraumatic Stress we were able to determine both predictive classification models for childhood PTSD and identify several causal variables. This set of techniques has great potential for enhancing the methodological toolkit in the field and future studies should seek to replicate, refine, and extend the results produced in this study. Keywords: Traumatic stress, PTSD, Machine learning, Informatics, Child & Adolescent psychiatry [7].

H. Predicting Insulin Resistance in Children Using a Machine-Learning-Based Clinical Decision Support System

Hall, A. J., et al., (2016, November) proposes a new diagnostic approach based on application of machine learning techniques to anthropometric patient features in order to create a predictive model capable of diagnosing insulin resistance (HOMA-IR). As part of the study, a dataset was built using existing pediatric patient data containing subjects with and without insulin resistance. A novel machine learning model was then developed to predict the presence of insulin resistance based on dependent biometric variables with an optimal level of accuracy. This model is made publicly available through the implementation of a clinical decision support system (CDSS) prototype. The model classifies insulin resistant individuals with 81% accuracy and 75% of individuals without insulin resistance. This gives an overall accuracy of 78%. The user testing feedback for the CDSS is largely positive. The current methods used by clinicians to identify insulin resistance in children are limited by invasive and clinically expensive blood testing. The benefits of this model would be to reduce the cost of clinical diagnosis and as a result, could also be used as a screening tool in the general childhood population [8].

I. Early 'Dengue Severity' Prediction

Caicedo-Torres, W., et al., (2016, November) advises that an Infection by dengue-virus is prevalent and a public health issue in tropical countries worldwide. Also, in developing nations, child populations remain at risk of adverse events following an infection by dengue virus, as the necessary care is not always accessible, or health professionals are without means to cheaply and reliably predict how likely is for a patient to experience severe Dengue. Here, they propose a classification model based on Machine Learning techniques, which predicts whether or not a pediatric patient will be admitted into the pediatric

Intensive Care Unit, as a proxy for Dengue severity. Different Machine Learning techniques were trained and validated using Stratified 5-Fold Cross-Validation, and the best model was evaluated on a disjoint test set. Cross-Validation results showed an SVM with Gaussian Kernel outperformed the other models considered, with an 0.81 Receiver Operating Characteristic Area Under the Curve (ROC AUC) score. Subsequent results over the test set showed a 0.75 ROC AUC score. Validation and test results are promising and support further research and development.

Applications of Machine Learning techniques in general health Care:

J. 'Depression detection' from social network data using machine learning technique

Islam, M. R., et al., (2018) recommends that Social networks have been developed as a great point for its users to communicate with their interested friends and share their opinions, photos, and videos reflecting their moods, feelings and sentiments. This creates an opportunity to analyze social network data for user's feelings and sentiments to investigate their moods and attitudes when they are communicating via these online tools. They aim to perform depression analysis on Facebook data collected from an online public source. They have evaluated the efficiency of their proposed method using a set of various psycholinguistic features. This proposed method can significantly improve the accuracy and classification error rate. In addition, the result shows that in different experiments Decision Tree (DT) gives the highest accuracy than other ML approaches to find the depression [10].

K. On the Prediction of Clusters for Adverse Reactions and Allergies on Antibiotics for Children

Yildirim, P., (2013, September) discussed that Antibiotics are the most commonly prescribed drugs in children and most likely to be associated with adverse reactions. Record on adverse reactions and allergy from antibiotics considerably affect the prescription choices. We consider this a biomedical decision problem and explore hidden knowledge in survey results on data extracted from the health records of children. We apply the K-means algorithm to the data in order to generate clusters and evaluate the results. As a result, some antibiotics form their own clusters. Consequently, medical professionals can investigate these clusters, thus gaining useful knowledge and insight into this data for their clinical studies [11].

L. A Comparative Study on Machine Learning Techniques for Prediction of Success of Dental Implants

Oliveira, A. L. I., et al., (2005, November) presents a comparative study on machine learning techniques for prediction of success of dental implants. The techniques compared here are: (a) constructive RBF neural networks (RBF-DDA), (b) support vector machines (SVM), (c) k nearest neighbors (k-NN), and (d) a recently proposed technique, called NNSRM, which is based on k-NN and the principle of structural risk minimization. We present a number of simulations using real-world data. The simulations were carried out using 10-fold cross-validation and the results show that the methods achieve comparable performance, yet NNSRM and RBF-DDA produced smaller classifiers [12].

M. Performance evaluation of different machine learning techniques for prediction of heart disease

Dwivedi, A. K. (2018) discussed that the heart diseases are of notable public health disquiet worldwide. Heart patients are growing speedily owing to deficient health awareness and bad consumption lifestyles. Therefore, it is essential to have a framework that can effectually recognize the prevalence of heart disease in thousands of samples instantaneously. At this juncture, the potential of six machine learning techniques was evaluated for prediction of heart disease. The recital of these methods was assessed on eight diverse classification performance indices. In addition, these methods were assessed on receiver operative characteristic curve. The highest classification accuracy of 85 % was reported using logistic regression with sensitivity and specificity of 89 and 81 %, respectively [13].

N. A machine learning approach for prediction of pregnancy outcome following IVF treatment

Hassan, M. R., **et al., (2018)** deals with a machine learning approach for prediction of pregnancy outcome following IVF treatment. Infertility affects one out of seven couples around the world. Therefore, the best possible management of the in vitro fertilization (IVF) treatment and patient advice is crucial for both patients and medical practitioners. The ultimate concern of the patients is the success of an IVF procedure, which depends on a number of influencing attributes. Without any automated tool, it is hard for the practitioners to assess any influencing trend of the attributes and factors that might lead to a successful IVF pregnancy. This paper proposes a hill climbing feature (attribute) selection algorithm coupled with automated classification using machine learning techniques with the aim to analyze and predict IVF pregnancy in greater accuracy. Using 25 attributes, we assessed the prediction ability of IVF pregnancy success for five different machine learning models, namely multilayer perceptron (MLP), support vector machines (SVM), C4.5, classification and regression trees (CART) and random forest (RF). The prediction ability was measured in terms of widely used performance metrics, namely accuracy rate, F-measure and AUC. Feature selection

algorithm reduced the number of most influential attributes to nineteen for MLP, sixteen for RF, seventeen for SVM, twelve for C4.5 and eight for CART. Overall, the most influential attributes identified are: 'age', 'indication' of fertility factor, 'Antral Follicle Counts (AFC)', 'NbreM2', 'method of sperm collection', 'Chamotte', 'Fertilization rate in vitro', 'Follicles on day 14' and 'Embryo transfer day.' The machine learning models trained with the selected set of features significantly improved the prediction accuracy of IVF pregnancy success to a level considerably higher than those reported in the current literature [14].

O. Evaluation of Machine Learning Techniques in Predicting Acute Coronary Syndrome(ACS) Outcome

Jaafar, J., et al., (2013) evaluates different machine learning techniques to present the potential use of machine learning techniques in classification tasks as basis for future medical prediction model development. A dataset of 960 of ACS patients from the Malaysian National Cardiovascular Disease Database registry was employed and trained on three popular machine learning classifiers i.e. Naïve Bayes, Decision Tree and Neural Network to predict ACS outcome. The outcome being evaluated was whether the patient is dead or alive. An open source tool—Waikato Environment for Knowledge Analysis (WEKA) were used in executing these classification tasks. A 10-folds cross validation technique was used to evaluate the models. The performance of classifiers was presented by their accuracy rate, confusion matrix and area under the receiver operating characteristic curve (AUC). Naïve Bayes and Neural Network show generally convincing results with an average of 0.8 AUC values and 90 % accuracy rate [15].

P. Using inductive machine learning, expert systems and case based reasoning to predict preterm delivery in pregnant women

Dyne, M., et al., (1994, September) extends a previously constructed prototype expert system to include case-based reasoning and learning, in order to improve the system's predictive accuracy in assessing preterm delivery risk. The initial expert system was developed by using an inductive machine learning technique on 9,445 data records of pregnant women, providing production rules to predict preterm delivery. Its predictive accuracy was tested on a separate set of 9,445 data records. Next, the capability to reason from both production rules and input test cases was added to the system, in addition to the capability to internally modify its confidence in each piece of knowledge (rule or case) and the relative importance of patient attributes which appear to be predictive of preterm delivery. The system was structured such that the accuracy of either type of reasoning could be measured individually to determine how rule-based and case-based reasoning perform alone, and to determine how they perform together. Results show that the predictive accuracy of the system was improved, with different trends emerging, dependent on the bias of the learning data, with the hybrid system providing the best predictive accuracy [16].

Q. Early Diabetes Prediction Using Voting Based Ensemble Learning.

Husain, A., et al (2018, April) have investigated the discriminative performance of ensemble learning model for diabetes prediction at an early stage. We have used different machine learning models and then ensemble it to improve the overall prediction accuracy. The dataset used is NHANES 2013-14 comprising of 10,172 samples and 54 feature variables for diabetes section. The feature variables used are in the form of questionnaire, a set of questions suggested by NHANES (National Health and Nutrition Examination Survey). An Ensemble model using majority voting technique was developed by combining the unweighted prediction probabilities of different machine learning models. Also, the model is evaluated and validated for real user input data for user friendliness. The overall performance was improved by Ensemble Model and had an AUC (Area under Curve) of 0.75 indicating high performance [17].

R. Machine Learning Solutions in Computer-Aided Medical Diagnosis(CAMD)

Belciug, S. (2016) achieves an Automated medical diagnosis by building a model of a certain disease under surveillance and comparing it with the real time physiological measurements taken from the patient. If this practice is carried out on a regular basis, potential risky medical conditions can be detected at an early stage, thus making the process of fighting the disease much easier. With CAMD, physicians can trustfully use the "second opinion" of the 'digital assistant' and make the final optimum decision. In this paper, they provide an overview on selected ML algorithms that can be applied in CAMD, focusing on the enhancement of neural networks (NNs) by hybridization, partially connectivity, and alternative learning paradigms [18].

S. A Machine Learning Approach to Prediction of Exacerbations of Chronic Obstructive Pulmonary Disease

Fernandez-Granero, M. A., (2015 June) deals with Chronic Obstructive Pulmonary Disease (COPD) places an enormous burden on the health care systems and causes diminished health related quality of life. The highest proportion of human and economic cost is associated to admissions for acute exacerbation of respiratory symptoms. The remote monitoring of COPD patients with the view of early detection of acute exacerbation of COPD (AECOPD) is

one of the goals of the respiratory community. In this study, machine learning was used to develop predictive models. Models robustness to exacerbation definition was analyzed. A non-knowledge based approach was followed on data self-reported by patients using a multimodal tool during a remote monitoring 6 months' trial. Comparison of different classifier algorithms operating with different AECOPD definitions was performed. Significant results were obtained for AECOPD prediction, regardless of the definition of exacerbation used. Best accuracy was achieved using a PNN classifier independently of the selected AECOPD definition. Our study suggests that the proposed data-driven methodology could help to design reliable predictive algorithms aimed to predict COPD exacerbations and therefore could provide support both to physicians and patients [19].

T. Machine Learning Method to Establish the Connection Between Age Related Macular Degeneration and Some Genetic Variations

Martinez-Velasco, A., et al., (2016) discusses that Medicine research based in machine learning methods allows the improvement of diagnosis in complex diseases. Age related Macular Degeneration (AMD) is one of them. AMD is the leading cause of blindness in the world. It causes the 8.7 % of blind people. A set of case and controls study could be developed by machine-learning methods to find the relation between Single Nucleotide Polymorphisms (SNPs) SNP_A, SNP_B, SNP_C and AMD. In this paper we present a machine-learning based analysis to determine the relation of three single nucleotide SNPs and the AMD disease. The SNPs SNP_B, SNP_C remained in the top four relevant features with ophthalmologic surgeries and bilateral cataract. We aim also to determine the best set of features for the classification process [20].

U. Dementia Screening with Machine Learning Methods

Shankle, W. R., et al., (1997) applied Machine learning algorithms to an electronic patient database generated by the UC Irvine Alzheimer's clinic to learn the simplest and most accurate patient parameters that would discriminate 244 very mildly demented from 198 normally aging subjects. Attributes included age, sex and education plus responses to the Functional Activities Questionnaire, the Mini-Mental Status and Blessed Orientation, Memory and Concentration tests. The machine learning algorithms included decision tree learners (C4.5, CART), rule inducers (C4.5Rules, FOCL) and naive Bayes. Stepwise logistic regression was used to compare results. The sample was randomly split into training and testing sets, and the results were validated over 30 runs. Although the Functional Activities Questionnaire has been used since 1980, the machine learning algorithms with access to the first trial identify that a single attribute, measuring forgetful-ness, equaled or exceeded the accuracy of any other scoring method of this test. Post hoc inspection of the odds ratios obtained by stepwise logistic regression confirmed this finding. The application of machine learning has identified an extremely simple, yet accurate screen for very mild dementia [21].

Table 1: Results

V. A Survey for identifying Parkinson's disease by Binary Bat Algorithm

S. Jabeen Begum & B. Swaathi2 (April 2019) described **t**hat Parkinson's disease is a chronic neurological disorder that directly affects human gait. It leads to slowness of movement, causes muscle rigidity tremors. Analyzing human gait serves to be useful in studies aiming at early recognition of the disease. In their present work, they perform a comparative analysis of various nature-inspired algorithms to select optimal features/variables required for aiding in the classification of affected patients from the rest. Binary Bat Algorithm (BBA) has searched the feature space for optimal feature combinations with good accuracy. Bats use echolocation to detect prey and avoid obstacles in the dark by emitting ultrasound waves and listening to the echo produced through the waves reflecting from the surrounding objects. The accuracy can be made possible through various classification techniques such as SVM, NB, k-NN [22].

III. METHODOLOGY

Children are vital assets of society. The tracking of data on children's health and its influences is an essential part of efforts to improve children's health and the health of the adults they will become. This paper aims to compare machine learning techniques that can predict various health problems of Children based on 22 papers accessed from the following electronic databases: IEEE Explorer, Science Direct, Springer, and Wiley.

This study has been carried out with the following steps:

1. To collect Children's various health issues that degrade their lift skill.
2. To collect the papers which predict their health issues using Machine Learning techniques.
3. To Compare the results of papers.
4. To analyze the accuracy of algorithms.
5. To list down the best algorithm to specific health problem.

IV. RESULTS AND DISCUSSION

Based on above research papers, the Table 1 represents the list of Health problems and the best algorithms to predict them.

Authors & Year	Health Problems	Best classification Model for Prediction
Dugan, T. M., et al (2015)	Early Childhood Obesity	ID3 model
Theodore, K., et al., (2018)	Anxiety Disorders	SVM
Shah, A. (2015)	Attention Deficit Hyperactive Disorder	Logistic Regression
Sumathi, M. R., & Poorna, B. (2016) and Dwivedi, A. K. (2018)	Mental Health Problems, Heart Disease	Multilayer Perceptron, Multiclass Classifier
Usta, M. B., et al (2018) and Hassan, M. R., **,et al.,(2018)**	Autism Spectrum Disorder, pregnancy outcome following IVF treatment	Naive Bayes and Neural Networks
Saxe, G. N., et al., (2017)	Child Post Traumatic Stress	Clinical decision support system
Hall, A. J., et al., (2016, November)	Insulin Resistance in Children	
Caicedo-Torres, W., et al., (2016, November)	Dengue Severity	SVM with Gaussian Kernel
Islam, M. R., et al., (2018)	Depression detection	Decision Tree

The data gathered from the survey represents that SVM produces best accuracy on anxiety disorders and dengue severity prediction. But it does not give better solution to other problems. While Naïve Bayes and Neural network algorithms provide best accuracy on Autism Disorder and Outcome of IVF treatment, they don't give best solution to other Problems. From this study we can conclude that each and every algorithm is unique on its result. Researchers should decide which algorithm is better to use depending on the health issues they have considered and the input dataset collected.

V. CONCLUSION AND FUTURE WORK

In this paper, we performed a survey on various machine learning techniques by classifying major algorithms, describing main ideas behind each algorithms and comparing their accuracy and reliability. The future is in the hands of the children who are going to rule the world with the invention in technologies. Early screening of their health problems could increase the chances of a healthy life for all children. With the help of Machine Learning techniques, we are paving the way to the future healthy world. This survey paper will help the researchers to compare and enhance their ideas of prediction.

In future work, an improvement regarding to which method is better to use depending on the type of health issues and how each type of health problems affects each method, is worth investigating. Also, by experimenting with other methods and approaches, we may achieve even higher scores for the prediction of health problems that could lead in a better treatment support for patients

.

REFERENCES

[1]. Dugan, T. M., Mukhopadhyay, S., Carroll, A., & Downs, S. (2015). Machine learning techniques for prediction of early childhood obesity. *Applied clinical informatics*, *6*(03), 506-520.

[2]. Theodore, K., Ioannis, L., Panagiotis, P., & Emmanouil, P. (2018). *Reviewing machine learning techniques for predicting anxiety disorders*.

[3]. Shah, A. (2015). Predicting 'Attention Deficit Hyperactive Disorder 'using large scale child data set.

[4]. Sumathi, M. R., & Poorna, B. (2016). Prediction of mental health problems among children using machine learning techniques. *International Journal of Advanced Computer Science and Applications*, *7*(1), 552-557.

[5]. Usta, M. B., Karabekiroglu, K., Sahin, B., Aydin, M., Bozkurt, A., Karaosman, T., ... & Sahin, İ. (2018). Use of machine learning methods in prediction of short-term outcome in autism spectrum disorders. *Psychiatry and Clinical Psychopharmacology*, 1-6.

[6]. Alarifi, H. S., & Young, G. S. (2018). Using Multiple Machine Learning Algorithms to Predict Autism in Children. In *Proceedings on the International Conference on Artificial Intelligence (ICAI)* (pp. 464-467). The Steering Committee of the World Congress in Computer Science, Computer Engineering and Applied Computing (World Comp).

[7]. Saxe, G. N., Ma, S., Ren, J., & Aliferis, C. (2017). Machine learning methods to predict child posttraumatic stress: a proof of concept study. *BMC psychiatry*, *17*(1), 223.

[8]. Hall, A. J., Hussain, A., & Shaikh, M. G. (2016, November). Predicting insulin resistance in children using a machine-learning-based clinical decision support system. In *International Conference on Brain Inspired Cognitive Systems* (pp. 274-283). Springer, Cham.

[9]. Caicedo-Torres, W., Paternina, Á., & Pinzón, H. (2016, November). Machine learning models for early dengue severity prediction. In *Ibero-American Conference on Artificial Intelligence* (pp. 247-258). Springer, Cham.

[10]. Islam, M. R., Kabir, M. A., Ahmed, A., Kamal, A. R. M., Wang, H., & Ulhaq, A. (2018). Depression detection from social network data using machine learning techniques. *Health information science and systems*, *6*(1), 8.

[11]. Yildirim, P., Majnarić, L., Ekmekci, O. I., & Holzinger, A. (2013, September). On the prediction of clusters for adverse reactions and allergies on antibiotics for children to improve biomedical decision making. In *International Conference on Availability, Reliability, and Security* (pp. 431-445). Springer, Berlin, Heidelberg.

[12]. Oliveira, A. L. I., Baldisserotto, C., & Baldisserotto, J. (2005, November). A comparative study on machine learning techniques for prediction of success of dental implants. In *Mexican International Conference on Artificial Intelligence* (pp. 939-948). Springer, Berlin, Heidelberg.

[13]. Dwivedi, A. K. (2018). Performance evaluation of different machine learning techniques for prediction of heart disease. *Neural Computing and Applications*, 1-9.
[14]. Hassan, M. R., Al-Insaif, S., Hossain, M. I., & Kamruzzaman, J. A machine learning approach for prediction of pregnancy outcome following IVF treatment. *Neural Computing and Applications*, 1-15.
[15]. Jaafar, J., Atwell, E., Johnson, O., Clamp, S., & Ahmad, W. A. W. (2013). Evaluation of Machine Learning Techniques in Predicting Acute Coronary Syndrome Outcome. In *Research and Development in Intelligent Systems XXX* (pp. 321-333). Springer, Cham.
[16]. Dyne, M., Tsatsoulis, C., & Thorp, J. (1994, September). Using inductive machine learning, expert systems and case based reasoning to predict preterm delivery in pregnant women. In *Database and Expert Systems Applications*.
[17]. Husain, A., & Khan, M. H. (2018, April). Early Diabetes Prediction Using Voting Based Ensemble Learning. In *International Conference on Advances in Computing and Data Sciences* (pp. 95-103). Springer, Singapore.
[18]. Belciug, S. (2016). Machine learning solutions in computer-aided medical diagnosis. In *Machine Learning for Health Informatics* (pp. 289-302). Springer, Cham.
[19]. Fernandez-Granero, M. A., Sanchez-Morillo, D., Lopez-Gordo, M. A., & Leon, A. (2015, June). A machine learning approach to prediction of exacerbations of chronic obstructive pulmonary disease. In *International Work-Conference on the Interplay Between Natural and Artificial Computation* (pp. 305-311). Springer, Cham.
[20]. Martínez-Velasco, A., Zenteno, J. C., Martínez-Villaseñor, L., Miralles-Pechúan, L., Pérez-Ortiz, A., & Estrada-Mena, F. J. (2016). Machine learning method to establish the connection between age related macular degeneration and some genetic variations. In *Ubiquitous Computing and Ambient Intelligence* (pp. 28-39). Springer, Cham.
[21]. Shankle, W. R., Mani, S., Pazzani, M. J., & Smyth, P. (1997). Dementia screening with machine learning methods. In *Intelligent Data Analysis in Medicine and Pharmacology* (pp. 149-165). Springer, Boston, MA.
[22]. S. JabeenBegum., B. Swaathi April (2019). A Survey for identifying Parkinson's disease by Binary Bat Algorithm in *International Journal of Scientific Research in Computer Science and Engineering* Vol.7, Issue.2, pp.17-23,

Authors Profile

Mrs. A.Meharaj Begum has completed Bachelor of Computer Science in 1999, Master of Computer Applications in 2002 and Master of Philosophy in 2005 from Manonmaniam Sundaranar University, Tamil Nadu, India. She has qualified SET in 2016 from Mother Teresa University for Women, Tamil Nadu and NET in 2018 from University Grant Commission, New Delhi, India. She is currently working as an Assistant Professor in Department of Computer Science, AIMAN college of arts and science for women, Bharathidasan University, Tiruchirapalli, TamilNadu, India. Her fields of interest are Machine Learning, Deep Learning Algorithms, Big Data Analytics and Data Mining.

