


A Review Study on Ensemble Machine Learning Approaches for Predication of Multiple Diseases and Causes for Women

Mr.S.A.J.Mohammed Haji Ali 
University of Technology & Applied Science - Ibra
Sultanate of Oman

Mr.V.Saravanakumar 
Assistant Professor,
Sri Venkateswara College of Engineering & Technology, Chittoor, Andhra Pradesh, India.
visaranams@gmail.com

Abstract

India today are organizing and providing subsidies for the empowerment of women to expose their changes in form of finance, health, education and family. To achieve that Government of India, Ministry of Human Resource Development (MHRD), Ministry of Women & Child Development and Non-profit organizations are providing awareness, welfare schemes, training programme and financial support via bank. Even though the growth of digital health care system, many people throughout world especially women facing lot of challenges in predicting and diagonalising the health issues. As per World Health Organization (WHO) 2021 statement that women are suffered with top diseases with high risk factors such as Mental Stress, metabolic Syndromes, Cardio Vascular Disease (CVD), Severe Complications during Pregnancy, Malnutrition and non-communication diseases. This paper works to bring out all recent methodologies, framework, protocols and approaches to early predication of many diseases in simple novel approach. Ensemble Machine learning approaches uses the multiple similar and independent techniques or models to derive the solutions from early predicting of Women's diseases with single instance and reporting the diagnosis and prognosis to reduce the cause of effect in continuation observation through machine learned or well trained equipments. As of report that ensemble ML methods with the best results were the prediction of premature diseases from medical images with an accuracy of 90.1-95.7%, and with XGBoost techniques with 94.4-98.7% accuracy.

Index Terms : Women, AI, ML, WHO, Ensemble

[I]. Introduction

Machine Learning (ML) technologies and its algorithms suited for dramatic changes in sense of converting unstructured data patterns such as images, videos, animation and soon into useful structured pattern with mathematical inferences of probability and statistics approaches[1,3]. Specially, machine learning approaches could support the medical practitioner and experts to improve health and closely observing all health related risk parameters. These ML algorithms applied in varies applications including medical image processing, automated text classification, visual differentiation, network intruder avoidance, reduction of junk email pushing, optimized credit card fraudulent and Disease optimizations[2-6].

Machine learning can be used in various ways to improve women's healthcare. Here are some examples[3]: Personalized Treatment[4]: Machine learning algorithms can analyze large amounts of data, such as electronic health records and genetic information[6], to develop personalized treatment plans for women based on their specific health conditions. Early Detection: Machine learning can help detect diseases like breast cancer at an early stage, by analyzing mammograms and identifying patterns that may indicate the presence of cancer. Fertility Treatment[7]: Machine learning can help fertility

clinics identify the best treatment options for women struggling with infertility, by analyzing patient data and predicting the likelihood of success for different treatments. Predictive Analytics[21,24]: Machine learning can help predict potential health risks for women by analyzing factors such as age, lifestyle, and family history. Disease Management: Machine learning can be used to manage chronic conditions such as diabetes, heart disease, and hypertension in women by analyzing patient data and predicting disease progression.

Most of the mentioned applications utilized the strategy of supervised learning approaches rather than unsupervised models. Supervised Learning: This type of machine learning involves training a model using labeled data to make predictions about new, unlabeled data. For example, supervised learning algorithms can be used to predict breast cancer risk based on a woman's medical history and family history. Unsupervised Learning: This type of machine learning involves training a model on unlabeled data to identify patterns or groupings within the data. Unsupervised learning algorithms can be used to cluster patients based on their medical histories or to identify common symptoms associated with certain health conditions. Reinforcement Learning[4]: This type of machine learning involves training a model to make decisions based on rewards or punishments. For example, reinforcement learning algorithms can be used to develop personalized treatment plans for women with chronic conditions such as diabetes or hypertension. In addition to these three main types of machine learning, there are also other specialized techniques such as deep learning, natural language processing (NLP), and computer vision that can be used in women's healthcare[20]. These techniques can be applied in various applications such as analyzing medical images, electronic health records, or patient-generated data from wearable devices. In supervised approaches there are several algorithms are classified under the most researchers outcomes as Naive Bayes algorithm, Support Vector Machine (SVM), Random Decision Forest, Linear Regression, Logistic Regression and Generalized Regression. The aim of supervised learning approach is to develop a prediction model with labelled (known) datasets and unlabelled (unknown) datasets[19].

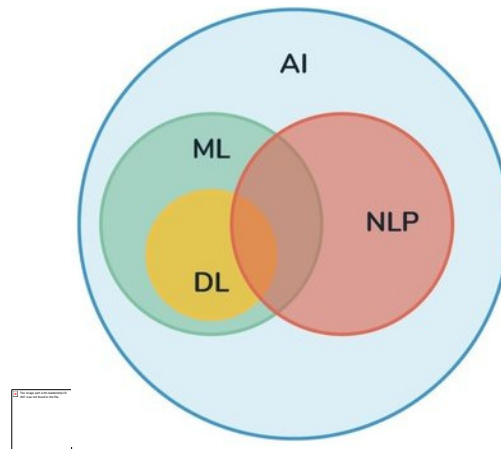


Figure 1.1 Relationship among AI,ML,DL and NLP

The primary objective of our research concentrate on developing a well labelled predication model for maximum diseases influenced among women through communicable and non-communicable diseases in early stages itself. There are many studies related to severe diseases that specifically affect women. Here are some examples: Breast Cancer[2,6]: There are numerous studies related to breast cancer, which is one of the most common cancers among women. Research in this area includes studies on risk factors, treatment options, and outcomes. Ovarian Cancer: Ovarian cancer is another serious disease that disproportionately affects women. Studies in this area include research on risk factors, screening methods, and treatment options. Cardiovascular Disease: Although often thought of as a male disease, cardiovascular disease is a leading cause of death among women. Studies in this area include research on risk factors, prevention strategies, and treatment options for women. Maternal Health: Maternal

health is an area of research focused on the health and well-being of pregnant women and new mothers. Studies in this area include research on prenatal care, pregnancy complications, and postpartum depression. Autoimmune Diseases: Autoimmune diseases such as lupus and rheumatoid arthritis are more common in women than in men. Studies in this area include research on risk factors, diagnosis, and treatment options for women. Overall, these studies help improve our understanding of these diseases and develop effective treatment and prevention strategies for women.

A novel framework illustrated which uses the backbone as neural network with various input sources such as datasets, medical images, questionnaires, survey analysis, speech and vocal systems. This novel framework All in One Women Prediction (AOWP) diseases in prognosis stages to avoid life long sufferance. Women are mainly affected with diseases after their mature and pregnancy life. The top diseases which is affected by women highlighted by World Health Organization (WHO) are Abnormal Blood Bleeding, Breast Cancer, Cardiac Vascular Diseases, Maternal Health issues, Mental (Brain neuron) imbalance and Malnutrition[8,9]. Several researches mainly focuses on delivering optimized electronic gadgets, tools, approaches, technologies and medicine to overcome diseases in post mature stages. This lead to loss of life and hope among the patient's. To avoid this conflict our framework will produce a optimal solution for predicating multiple diseases and disorder with single prognosis test.

In recent years, deep learning that is a part of machine learning opens the door for Biometric diseases diagnosis. As per report of WHO, pregnancy associated death is every day 808 women (2017) due to malnutrition, improper observation during the stages of pregnancy, fatigue, geographical function and careless of proper medication. Another 1/4th of global population are affected with Autism Spectrum Disorder (ASD) for women are less than men and Parkinson's Disease (PD). As per 2025 ASD might grown to 1 in 2 due to digital technology, sensor technology, physical less maternity and poison junk foods. Non-communicable Disease (NCD) is a recent hazardous disease that revolved past two decades with high risks and maximum 3/4th of women death of every year mostly in under developing and poorer economic countries. This DL records the time series data, health records and images (PET Scan,X-rays, MRI, CT Scan , Ultra Sound, ECG). PD's motor symptoms hypo phonic speech , tremor and sleep less disorder, hallucinations, brain impairment of non-motor's are resolved nowadays with the support Conventional Neural Network (CNN) and Recurrent Neural Network (RNN) approaches.Deep Learning based models provides rich variety of medical applications like health records, time series model and health records by discovering hidden informations from unstructured (images, videos) as an input, processed through CNN & RNN algorithm and finally producing target output visualization. Ensemble classifiers support many image processing data produce precision and accurate results with valid data set and training data set. Ensemble Learning uses stacking (improve predications), Sequential (reduce the bias) and Parallel (use decrease variance). Our model uses all the ensemble learning algorithms to produce well suited optimal supervised Disease predictor for all cases in one.

The rest of the paper is organized as follows[12,14]: Related study is discussed deeply in Section 2, Section 3 describes the methodology and approaches for existing system, novel method and architecture were described in section 4 and finally conclusion in section 5.

[II]. Related Study

In medical field, all the diseases for woman were categories into four types namely infection, deficiency, genetic and non-genetic diseases. Some of the most specific health related diseases affected by female with high risks were diabetes, stoke, heart, mental health issues, uterine issues, breast cancer and osteoporosis. This section will elaborate the symptoms, impact factor, diagnosis and electronic equipments that leads the life of female from early age to old ages in detail.

Abnormal Uterine Blood Bleeding(AUB) or High Menstrual Bleeding (HMB) is one of the immature and inherited disorders commonly among the 15-49 ages of woman. Abnormal Uterine Bleeding (AUB) is a common gynecological problem that affects women of all ages. Deep learning, a type of machine

learning that uses artificial neural networks, has the potential to improve the accuracy of AUB diagnosis and treatment. One study published in the Journal of Medical Systems[6,29] used deep learning to classify AUB cases based on clinical data and sonographic features. The researchers used a convolutional neural network (CNN) to analyze ultrasound images and extract features related to AUB. They also used a long short-term memory (LSTM) network to analyze clinical data such as patient age, menstrual cycle history, and hormone levels. The results showed that the deep learning model was able to accurately classify AUB cases with a high degree of accuracy. The study concluded that deep learning has the potential to improve the accuracy of AUB diagnosis and treatment by analyzing large amounts of data and identifying patterns that may be difficult for human clinicians to detect. Another study published in the Journal of Ultrasound in Medicine used deep learning to predict the need for hysteroscopy in women with AUB. The researchers used a CNN to analyze ultrasound images and extract features related to AUB, and then used a logistic regression model to predict the need for hysteroscopy. The results showed that the deep learning model was able to accurately predict the need for hysteroscopy in women with AUB, with a sensitivity of 96.9% and a specificity of 87.5%. The study concluded that deep learning has the potential to improve the efficiency and accuracy of AUB diagnosis and treatment by reducing the need for unnecessary hysteroscopy procedures. Menorrhagia is a heavy flow of blood during periods which may be prolonged more than a week. Since there are a lot of confusion arises in defining the terminology of AUB or HMB, even though the terms International Federation of Gynaecology and Obstetrics (IFGO) clears that both are associated with structural causes and non-structural causes. Due to heavy periods, the female may have possible diseases to be affected are fibroids, endometrial polyp, endometriosis and anticoagulant. During the therapy treatment of adolescents resulted that 20-70% of woman patient associated with Anticoagulation and increased from 17.8% before HMB. Kathryn E. Dickerson (2018) organized survey among experts of international physicians points that 10-62% of adults with HMB may be underlying Blood Disorder (BD).and also suggest that prognosis of AUB/HMB in early stage will support reduce the major risks in pregnancy, hormonal suppression, malnutrition, improve reproductive system. It is quite clear that failure to do early prediction could reduce the sufferance and risks of serious bleeding stage. AUB in early mature stage may leads to affect 9-14% of women in reproductive years. Girija D.K (2013) et.al developed a special neural based network system called Fibroid Disease Prediction System (FDPS) to determined most related fibroid diseases with ten risky parameters such as age, bleeding condition (low, moderate or heavy), marital status, pelvic pain results that 98% accuracy were obtained. Lola Gomez Jemes (2022) build a ML system to analysis and forecast intrauterine growth rate which mainly restricts pregnancy period. They compared many ML models and sorted based on the accuracy of results. Accuracy of most used and supported models are Extra trees with 0.7894, Binary relevance—Random Forest 0.7368, Binary Relevance—Gaussian NB, SVC with 0.6315, K Neighbors Classifier with 0.5789, Binary Relevance—Decision Tree Classifier with 0.5263 and Dummy Classifier with 0.2105. Finally, they developed an ML model extra tree classifier to expose that this model had best metrics to evaluate all existing model in sense of AUC.

Next complications occur during pregnancy time and impacts were hypertension, preeclampsia, preterm labor, depression, anxiety and Miscarriage. During 2022 Vesna D. Garovic et.al gave a detail studies of Hypertensive Disorders of Pregnancy (HDP) . The research work on Machine Learning Models to Predict Hypertension in Women Using Clinical Data and Demographic Information found that both logistic regression and GBM models achieved high accuracy in predicting hypertension in women. The GBM model achieved the highest accuracy of 89%, with a sensitivity of 82%, specificity of 87%, and an area under the ROC curve of 0.92. Chronic hypertension is one of the major cause for maternal mortality in globe wise. As per their survey report that nearly 7.5% female were infected due to pregnancy oriented complications. So many treatments were implemented to reduce the high rated complication specially Systematic Vascular Resistance in order to decrease the volume of cardiac arrest, premature infants and genetic disorders. Munetoshi Akazawa (2021) et.al made a clinical tests with

sample of 9,894 pregnant women in Tokyo with median age of 31 published that 188 cases (1.9%) of women suffered heavy blood losses. They used technologies of deep learning and machine learning models with different features maternal weight before pregnancy, periodic weights, weight during labor admission, height, age, weight,height & sex of baby, and the fetal position. A Systematic review (2020) addresses 14 prediction models developed of postpartum hemorrhage (PPH) in vaginal cases. Their studies addresses that lack of features, predication model, algorithms and number of PPH cases were low results to lower predication accuracy. Thus they concluded that many big data cases to be evolved over the world and applying trendy methodology to develop models to have precision and accuracy of results. There are many tools and apps were developed for continuous monitoring of pregnancy woman for better identification of symptoms using chatbot called Velmio. This app generated detailed and elaborated studies that could be shared with their doctors which helps in clinical decision making.

Female were mostly affected types of cancer as per U.S medical center were skin, vaginal, uterine, breast, and colorectal. The symptoms such as over bleeding during periods, severe pain after sexual relationship, continuous pain in the lower abdomen more than two weeks and pain during urine flow were clears that female is affected by uterine cancer. Similarly the symptoms like painless breast, swelling, nipples turn inwards, colour changes or irritation around the nipple and discharge from the nipple were the primary causes for female breast cancer. Early detection of breast cancer is only key way to overcome the treatment options. With the invent and advent of Artificial Intelligence (AI) and recent Deep Learning techniques has been facilitate in effective and early diagnosis of breast cancer detection. Breast cancer happened in four ways namely: normal, benign, in-situ carcinoma and invasive carcinoma. A study published in the Journal of Digital Imaging used deep learning to classify breast lesions on ultrasound images. The researchers developed a convolutional neural network (CNN) model that achieved an accuracy of 95.9% in differentiating benign and malignant lesions. Another study published in the Journal of the American Medical Informatics Association used deep learning to predict breast cancer risk based on mammogram images. The researchers used a CNN model that achieved an area under the curve (AUC) of 0.73 in predicting breast cancer risk. Maged Nasser (2023) studies that Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) based on 98 articles summarized and concluded that Conventional Neural Network (CNN) is more accurate than conventional methods for breast cancer analysis. Several deep learning models has been evolved in past two decades in that foremost models included CNN, RBM and RNN, Deep Autoencoder, and GAN has proven in image analysis, computer vision, recommender system and NLP. They collected the article data from the year of 2010 to 2021that number of articles published for predication (Early Detection) of Breast Cancer with number of articles with journal.Compared to genetic data, imaging data are more readily available as datasets. They concluded that hybrid algorithms of CNN and booster algorithm increases the percentage of accuracy than existing models. According to the author Zhou X (2020) and his team, MRI, CT, US, X-Ray, and PET images supported to predict the cancer occurs. These breast type cancers are diagnosed mostly by : histopathological image analysis or genomics. As per the author review image noise are reduced by KNN, SVM, NN models and clearly initiated that Deep CNNs models would produce more precision and accuracy for early identification and treatments. They took survey from top databases IEEE, Elsevier, Nature, PubMed and Springer from 2010 to 2020. Most of the researchers and citations implied that DNN with hybrid algorithms produce remarkable precision than existing traditional technologies.

Demilade A. Adedinsewo (2022) et.al described the overview of key factors for cardiovascular disease(CVD) influencing women in traditional as obesity, hypertension, tobacco usage and traceable approaches were prescribed for CVD with rapid and new development of machine learning tool with the joint of artificial intelligence. As the resultant survey CVD cause of death ratio revealed that 64% in 2009 had been decline to 44% in 2019 due to several awareness activities. Journal of the American College of Cardiology used a convolutional neural network (CNN) to analyze cardiac magnetic

resonance images and accurately predict the presence of CAD in women. The model achieved an accuracy of 86.8%, which was higher than that achieved by human readers. Journal of the American College of Cardiology used a deep learning model to predict the risk of AF in women based on electrocardiogram (ECG) data. The model achieved a high accuracy of 85% in predicting the risk of AF in women. A study published in the Journal of the American Society of Echocardiography used a deep learning model to analyze echocardiography images and accurately diagnose VHD in women. The model achieved a high accuracy of 91.1%, which was higher than that achieved by human readers. Screening CVD in prevention stages 1. Primary Prevention done by awareness, counselling, education, lifestyle awareness and proper therapy consultation 2. Secondary Prevention attained with early diagnosis and prevent complications through continue medication, lifestyle therapy, monitoring and recurrent events with the help of AI tools. As per the quote of Geroge Box “ All models are wrong, but some are useful”. In the era of digital century AI combines statistical and regression techniques build tools that were most useful to prognosis and diagnosis CVD. ECG images of unstructured data were transformed to structured data that added more values for identification and assessment of any image processing diagnosis. A CNN were made researchers to delivered modern tools and techniques to model the Electro Cardiogram (ECG or EKG) that convert heart signals to meaningful pattern to detect several supplements of left ventricular dysfunction, hypertrophic obstructive cardiopulmonary, stiff heart syndrome, Aortic valve stenosis and normal sinus rhythm. The field of cardiology has tremendous growth with the support of AI & ML approaches and tools in sense of smart appliances such as Smart watches and wearable fitness trackers. These tools helps to identified risky people with 93 % of left ventricular dysfunction and 85% accurate in mammogram. Applying AI to ECGs, resulted that low cost test to easily identified weak heart valve and detect faulty atrial fibrillation before any symptoms were evidenced. Most of the experts in CVDs concentrated to develop a system with monitoring of arrhythmia and embolic stroke with high technology tools and smart gadgets.

Malnutrition, inadequate or imbalance nutrition results to hyper metabolism, inflammation, or malabsorption. Malnutrition may be under nutrition or over nutrition. Over nutrition might cause obesity or over weight that might leads to coronary artery diseases and stroke. health complications. Deep learning, a type of machine learning, has the potential to improve the detection, prevention, and treatment of malnutrition in women. Deep learning algorithms can analyze large amounts of data, including medical history, dietary habits, and demographic information, to identify patterns and predict malnutrition risk factors. This information can then be used to develop personalized nutritional plans and interventions for women at risk of malnutrition. One study published in the Journal of Medical Systems used deep learning to predict malnutrition risk in pregnant women based on clinical and demographic data. The researchers used a deep neural network to analyze data such as age, body mass index, dietary habits, and medical history. The model accurately predicted malnutrition risk in pregnant women, identifying several key risk factors that could be targeted for prevention. Another study published in the Journal of Nutrition used deep learning to develop personalized nutrition plans for pregnant women based on dietary habits, medical history, and other factors. The researchers used a deep neural network to analyze data from a large population-based study and develop a personalized nutrition plan for each participant. The model accurately predicted nutrient intake and identified specific dietary recommendations for each participant, improving their overall nutritional status. Most of the under developing countries suffers due to under nutrition. UNICEF[30] for every child reported quarter of women with less BMI 18.5 kg/m leads to birth to undernourished baby and inter cycle of nutrition growth of child is low. Malnutrition is a self-diagnosable to overcome the effects of weight loss, dizziness, fatigue, slow growth, thin muscles, anemia and water electrolyte imbalance. Periodic BMI checking and based on the results lapse of vitamins A or D should be increase through natural or medication. Another solution if continuous analysis supplements of micronutrients through blood test. Child Growth Monitor (CGM) is one of the most invited application that replaces all conventional strategy or approaches of anthropometric metrics which are complex, slow, and expensive, frequently

resulting in low weight data and improper assessments. CGM predicts measurement of height, weight, and mid-upper arm circumference (MUAC) of children below the age 5 with the art NN algorithms to predicate if a kids were under nourished or not. Boyang Tom Jin (2022) introduced model for Predicting malnutrition from longitudinal patient trajectories with deep learning. Early detection and diagnosis of imbalance nutrition had a major impact on clinical and economical outlets. Once woman has been identified with malnutrition can be reduced through deep learning models were developed with RNN architecture Long Short term Memory (LSTM). Since LSTM is a time series predicated data and easy to categories malnourished or not. An hybrid approach, using LSTM and XGBoost supports fine tuned predictive models. In the year of 2021 Haile Mekonnen Fenta reported that ML approaches supported to predicate main five under nutrition in Ethiopians. They collected all survey data of the year 2000, 2005, 2011 and 2016 categories six main ML algorithms with evaluation parameters in terms of sensitivity, accuracy, F1 score , precision and area. They used nearly 23467 training dataset and 5866 testing dataset to develop Six ML models to choose best evaluator. However RF model were chosen as an classifier engine to construct the malnutrition system for Ethiopian administrative areas.

A heart attacking is one of the serious life threatening event caused for men and women. The major 8 signs for possible heart attack[9] were chest pain, fainting, feeling light-headed, slow breathing, arm or joint pain, less sleeping, vomiting and less sleeping respective age, family history, medical follow-up, habits and food culture. Since CNN with hybrid booster algorithms helps to measure the cognitive and psychological stress faced among woman in any environment. 2/3rd of people are suffered regularly chronic stress might results in CVD, cancer, depression, and diabetes. Prior identification and analyse of physiological signal were much reliable to drive to next level[5,7]. Healey and Picard conducted one of the first studies that used physiological s ignals to detect the presence of human stress. The researchers used signals collected from the electrocardiogram, electromyography, electrodermal activity, and respiratory rate sensors. 22 features were hand-crafted from the aforementioned physiological signals. The LDA machine learning algorithm was used for binary classification between a stressed condition and a non-stressed condition. Russell Li (2020) addresses that 70 % Americans exercise stress due to the environment impact which raise risks of depression, post-traumatic stress disorder, violent victimization and loss of interest in living. According to their research, accuracy might decreases by choosing conflict attributes. They designed and demonstrated two models for binary and 3-class classification. The deep CNN achieved 99.80% and 99.55% whereas deep multilayer perceptron neuron network achieved only 99.65% and 98.38%. Thus their results clears that deep neural networks performed significant results than conventional machine learning methods.

[III]. Existing Approaches

With the era of Artificial Intelligence (AI) with Machine Learning (ML) illustrated more advances in medical technologies in the way of tools, equipment, mobile Apps which gives drastic changes in the health care industry. We remembered the words of George E.P. Box that “ All models are wrong, but some are useful”. With that statement, this chapter will discuss all the existing and useful ML algorithms in contrast to traditional software development which programs better outcomes of human tasks[18,22]. ML has been defined in many research area with edge evolving and broad techniques. Supervised learning, unsupervised learning and reinforcement learning are the primary types of machine learning. Machine learning methodology can be used to develop predictive models for various women's diseases[22-31]. The methodology typically involves the following steps:

1. Data collection: Collecting relevant data from electronic health records, clinical databases, or other sources.
2. Data preprocessing: Cleaning and preprocessing the collected data to remove missing values, outliers, and other data quality issues. Categorical variables are encoded using one-hot encoding, and continuous variables are standardized to have a mean of 0 and standard deviation of 1.

3. Feature selection: Identifying the most relevant features or variables that have the highest predictive power for the disease of interest. This is typically done using statistical methods or machine learning algorithms.
4. Model development: Developing machine learning models that can predict the occurrence or severity of the disease based on the selected features. Commonly used algorithms include logistic regression, decision trees, random forests, support vector machines (SVM), and neural networks.
5. Model evaluation: Evaluating the performance of the developed models using metrics such as accuracy, sensitivity, specificity, and area under the receiver operating characteristic (ROC) curve. Cross-validation is commonly used to ensure the models generalize well to new data.
6. Model deployment: Deploying the developed models in clinical practice to aid in diagnosis, risk assessment, and treatment planning for patients[1,8,3].

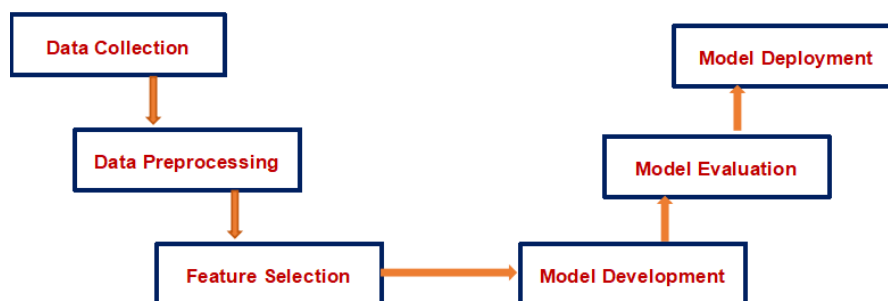


Figure 3.1: General Methodology

In supervised learning, labelled data is the prerequisite of predefined target variables and known values of feature variables[19,31]. Developing a model with a relationship between feature and victim attributes is well known as model processing. If the interested target attributes is continuous then it comes under the category of regression whereas attribute or variable is categorical then it is classification. Unlike supervised learning, unsupervised learning doesn't have any target variables, it is used to discovered useful pattern among the similarities and dissimilarities among objects. In this sense, unsupervised learning used to cluster the required units namely cluster analysis and typical interference relationship between customers to customers or object to objects or customers to objects or objects to customers is known as association analysis. While reinforcement learning applied to the situations in which reward function plays action based on action instead of correctness of responses. In context to business, reinforcement approach has been used commonly in business sector since they needs the profitableness than functionality. These three types of ML has several methods such as linear regression, logistic regression, decision trees, support vector machines, artificial neural networks and ensemble.

Linear Regression (LR) used to predict the future relationships among the variables or attributes. It may be further classified as dependent and independent variables. Depend on this, LR will be categories as simple linear regression and multiple linear regression. Naive Baye's algorithm works independently irrespective of their variable correlations. Most commonly used for text classification and expert recommender systems. The best approaches used in business organization is decision tree which has ID3 and CART to the labelling of forecast data. Classification is a unique type of supervised learning algorithm that relegate the data into various categories or classes. Some of the commonly used classification algorithms are K-nearest neighbor, Random forest, support vector machine and linear classifiers. Logistic regression estimates binary outcome either pass/fail, Yes/No falls on the basis of dependent and independent variables[23]. This method mainly used to determine whether the object in the image is presence or not by the value of 0 or 1. In medical research, better data analysis with either event occur or not, to be simple patient alive or not. Naive bayes in text analysis to calculates the

possibility of data point exist to certain class or not. Since Naive bayes is based on well known Bayes theorem where is a way of finding probability when we have a certain other probabilities. Past two decades used in sentimental analysis to differentiate the positive and negative reviews of patient's health system. These approach were well applied in the field of bioscience for predictive representation of various diseases (brain tumor, Bronchial asthma , prostatic carcinoma, and breast cancer and focuses on stages 0, I, II, and III). Multinomial distribution is an event model commonly used for document classification problems related to categories large volume of data items in the relevant database or inventories. Bernoulli Naive Bayes Classifier is one of the variant of naive bayes classifier in machine learning. It is useful to predict whether a person has a disease or not based on available data. A special type of NB algorithm is Gaussian Naïve Bayes Classifier, useful to estimate the likelihood of datasets where the data is to be continuous. K-nearest neighbour approach is a simple, very effective, high predictive, non parametric and pattern recognized algorithm that uses training datasets to the k closest relative features. Heart disease is one of the most dangerous and deadly cause for death in under developing countries specially in India. KNN and genetic algorithm produces many medical models to digest from severe heart, gastric cancer, tumor and etc.

The main aim of unsupervised learning is to predict underlying pattern or structures by grouping similar features and represent data sets in a clustered format. Based on the types of problems, classified as clustering and association rules approach. Apriori (2013) method is useful to identify frequent diseases from various geographical data with different time series. The combination pattern mining with apriori algorithm helpful in predict diseases in early stages, reduces death mortality and large diagnosis of human body system. Its proven that K-means clustering with naive bayes helps in prognosis and diagnosis of diseases either communicable on non-communicable. Bipolar neural disorder among women specially imbalance due to working and family communication exponentially increases in the past two decades[30]. Cluster based classification that supports for hospital management to allocate patients bed based on EHR severity. Since data collected from multi hospital is a big data and research contribute to handle big data research is a complex and huge step processes. Most of the clinical data derived from heterogeneous patient details are aimed to clustered as hierarchical[3]. This Hierarchical Cluster Analysis (HCA) divided as agglomerative (bottom-up) and divisive (top-down) approaches based on the way data's are splits and recursive. Recent research works focussed to visualized their results using heat map for different colours and that derives variety of controls on data. The most of clinical decision system tries to diagnosis and reduces the complexity of diseases among patients. Since high dimensional clinical data from meta world increases the complexity and reduces the percentage of prognosis and diagnosis. Thus building any type of model, its essential to identify the high complex attributes or features to extract quality data. As of now, dimension reduction and feature selection are two main methods used in dimensional reduction. Apart from these two, there are Principal Component Analysis (PCA), Non-negative matrix factorization (NMF), Linear discriminant analysis (LDA), Generalized discriminant analysis (GDA) and Missing Values Ratio also used for attribute feature selection and reduction with machine learning technologies. Breast tumor and kidney proteomics are determined and best parameters were obtained using classical technique called PCA and newer Diffusion Maps replaced with accurate and effective precision detection. The growth of Artificial Intelligence lays a path of another type of hierarchical or multi layer perception which loosely named similar to the human neuron structure is deep learning[9]. This deep neural network (DNN) uses different mathematical layers that translates raw data to meaningful outcomes and most of the layers used for conversion and refining the output. Medical Imaging is the vital input in clinical system, to analyse and predicate the images to diagnosis many retinal diseases. Sequential or temporal data are used in Recurrent Neural Network (RNN) to rectify the problems in NLP and image captioning. The hybrid model used both CNN and RNN for significant results. Generative Adversarial Networks (GAN) is most commonly used the inputs of image, video and voice generation. It is highly recommended to protect patient privacy. The deep ensemble machine learning

method in health care industry as produce tremendous results due to its two forms bagging and boosting for better performance in prediction and diagnosis. In addition, ensemble machine learning techniques generates strong models that gains more reliability when combining predictive and classification models[28].

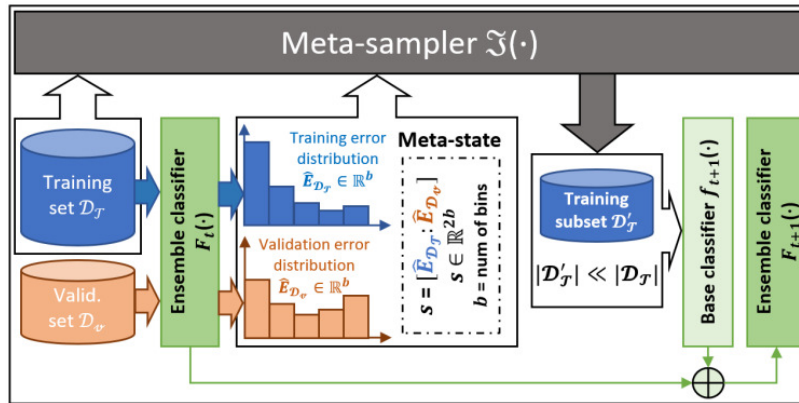


Figure 3.2: Ensemble Classifier (General)

Later the growth of deep learning, Ensemble associated with DL to predict many image map medical complex diseases. For instance, most of the brain, lungs, kidney and heart related issues were modelled with deep ensemble neural network. Many evidence shows that ensemble with deep learning had more advantages than traditional machine learning methods. For an example, two deep learning architecture namely inception V3 and ResNet 50 supports to classified different types of skin diseases. Alzheimer's diseases diagnosis had more growth rate with combination of sparse regression and deep learning methods. Most of the traditional ML approaches deployed with DNN specially random forest, stacking classifier, and gradient boosting which produces better results with hyperparameter.

[IV]. System Methodology

After Analysis of different types, methods and techniques of AI related contribution in health care industry. Its clear that most of the conventional approach produces less results than hybrid methods. Since many of the women diseases were unidentified to less significant tools, methods and technologies. AI markets focussed on benefits of business and a complete tool or components resolving the woman based health diagnosis is less. Hence its needs to combined components, tools, strategy and technologies to enhance the future of woman's better health system either in home, office and convenient areas[25]. This paper proposed a novel based framework that tries to reduce the complexity of identifying and diagnosis of health system. This framework which takes multiple inputs such as data sheets, videos, voices, electronic patient report, images etc and convert them to structured process data through the translator of deep neural networks with high dimension feature reduction and these data sets are undergone as inputs for ensemble learning system which is a combination of optimized bagging and booster. This system is collective data processing and producing target patients outcome for early detection of impact of existing or new communicable and non-communicable woman diseases. Designing a complex system in this digital era might results better, even though the solutions derived by this system could not matches with individual existing products. But today medical system needs to introduce an unique collective systems to easily collected all the hyper parameters that could reflects the future inconsistency. Here proposed novel design might helps the future researcher or industrial product developer to design a product which will early detect and diagnosis more than one affected issues.

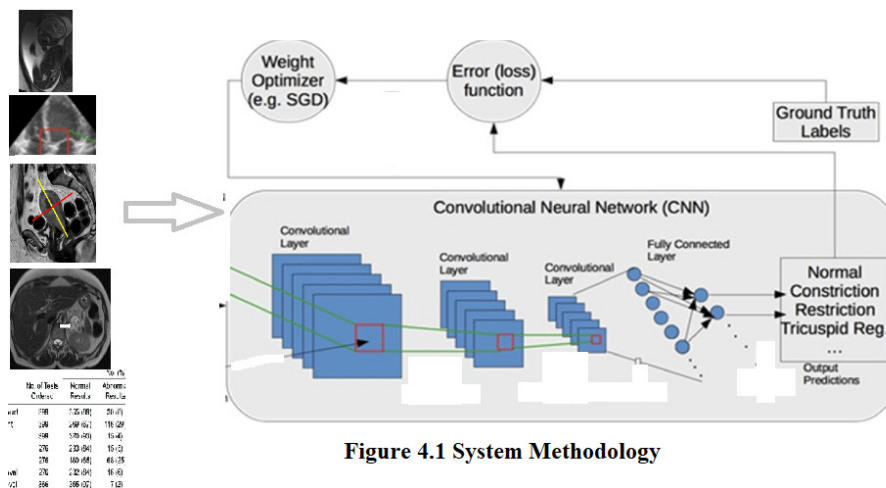


Figure 4.1 System Methodology

There are several drug or product makers thinks about to developed a single product or product which gives a results of multiple diseases data results. This will support the doctors to prevent or diagnosis diseases to create a zero disease environment. Cipla’s cippoint product used the wide range of testing parameters that includes fertility, inflammation, thyroid, diabetics[26], cardia and metabolic identification. Since the success rate is not appropriate to the market risks, but developing one unique product and continuous research and development will helps to disease free environment.VoCare platforms through mobile based technology managing multiple severe conditions in diabetes, CHF, CoPD and hypertension. Since the world is looking for device that should be cost less, portable, flexible, accurate and optimal. Hence its not a person or a sector delivering a wearable devices or portable equipment will meet all the constraints declared by WHO or FDA. With the global pandemic embassies among the world needs a innovative and modern framework that should continuously monitoring, diagnosing and treatment through electronic health care system (eHCS) proposed a system which might meets the need by integrating different framework in all levels of technology in Industrial 5.0.Figure 4.1 above shows that input from various data (blood, DNA,image, video and voice) , process (building multi layer network that processed high dimensional big data) and target output (expectation of medical people) to meet all in one.

Conclusion and Research Scope

The healthcare system for women in developing countries specially India has made significant progress over the years, but there is still much work to be done to ensure that women receive high-quality healthcare services. Despite the efforts made to improve maternal and child health, the maternal mortality rate in India remains high, and many women lack access to basic healthcare services. There are several factors contributing to the challenges faced by the healthcare system for women in under developing countries including limited resources, insufficient funding, and poor infrastructure. Cultural and social norms also play a significant role, with many women facing discrimination and barriers to accessing healthcare services. To improve women's healthcare, there needs to be a greater investment in healthcare infrastructure, training for healthcare providers, and increased funding for healthcare services. Efforts should also be made to address the cultural and social barriers that prevent women from accessing healthcare services, such as gender-based violence, poverty, and lack of education. Recent years in India, there have been several initiatives aimed at improving women's healthcare in India, including the National Health Mission and the Janani Suraksha Yojana, which provide financial assistance for maternal and child health services. The government has also launched various schemes to improve healthcare infrastructure and encourage private sector investment in healthcare. In conclusion, while progress has been made, there is still much work to be done to improve the healthcare system for women in India. Deep learning has the potential to revolutionize the diagnosis, treatment, and management of various women's diseases. There are several research areas where deep learning can be

applied to improve healthcare outcomes for women, and further research in these areas is needed to realize the full potential of this technology.

References

1. Jiang F, Jiang Y, Zhi H, et al. Artificial intelligence in healthcare: past, present and future. *Stroke Vasc Neurol* 2017;2:230
2. Becker A. Artificial intelligence in medicine: what is it doing for us today? *Health Policy Technol* 2019;8:198–205.
3. Harrell FE, Lee KL, Mark DB. Multivariable prognostic models: issues in developing models, evaluating assumptions and adequacy, and measuring and reducing errors. *Stat Med.* 1996; 15:361–387.
4. Guan Y, Nehl E, Pencea I, Condit CM, Escoffery C, Bellcross CA, McBride CM. Willingness to decrease mammogram frequency among women at low risk for hereditary breast cancer. *Sci Rep* . 2019;9(1):9599.
5. Burnside ES, Liu J, Wu Y, Onitilo AA, McCarty CA, Page CD, et al. Comparing Mammography Abnormality Features to Genetic Variants in the Prediction of Breast Cancer in Women Recommended for Breast Biopsy. *Acad Radiol* . 2016;23(1):62–9. doi: 10.1016/j.acra.2015.09.007.
6. American Cancer Society. *Cancer facts & figures 2018* . Atlanta: American Cancer Society; 2018.
7. Janghel RR, Shukla A, Tiwari R, Kala R. Breast cancer diagnosis using artificial neural network models. *The 3rd International Conference on Information Sciences and Interaction Sciences*; Chengdu, China: IEEE; 2010. p. 89-94 .
8. Stephens K. New Mammogram Measures of Breast Cancer Risk Could Revolutionize Screening. *AXIS Imaging News* . 2020.
9. Atashi A, Sohrabi S, Dadashi A. Applying two computational classification methods to predict the risk of breast cancer: A comparative study. *Multidisciplinary Cancer Investigation* . 2018;2(2):8–13. doi: 10.30699/acadpub.mci.2.2.8.
10. I. Sartzetakis, K. Christodoulopoulos, and E. Varvarigos, “Accurate quality of transmission estimation with machine learning,” *Journal of Optical Communications and Networking*, vol. 11, no. 3, pp. 140–150, 2019.
11. Verma D, Mishra N. Analysis and prediction of breast cancer and diabetes disease datasets using data mining classification techniques. *2017 International Conference on Intelligent Sustainable Systems (ICISS)*; Palladam, India: IEEE; 2017. p. 533-8 . doi: 10.1109/ISS1.2017.8389229.
12. Feld SI, Fan J, Yuan M, Wu Y, Woo KM, Alexandridis R, Burnside ES. Utility of Genetic Testing in Addition to Mammography for Determining Risk of Breast Cancer Depends on Patient Age. *AMIA Jt Summits Transl Sci Proc* . 2018;2017:81–90.
13. Blandin Knight S, Crosbie PA, Balata H, Chudziak J, Hussell T, Dive C. Progress and prospects of early detection in lung cancer. *Open Biol* . 2017;7(9):170070. 14. Jothi N, Husain W. Data mining in healthcare-a review. *Procedia Computer Science* . 2015;72:306–13.
14. Brédart A, Kop JL, Antoniou AC, Cunningham AP, De Pauw A, et al. Clinicians’ use of breast cancer risk assessment tools according to their perceived importance of breast cancer risk factors: an international survey. *J Community Genet* . 2019;10(1):61–71.
15. O. Stephen, M. Sain, U. J. Maduh, and D.-U. Jeong, “An efficient deep learning approach to pneumonia classification in healthcare,” *Journal of Healthcare Engineering*, vol. 2019, 2019.
16. Hou C, Zhong X, He P, Xu B, Diao S, Yi F, Zheng H, Li J. Predicting Breast Cancer in Chinese Women Using Machine Learning Techniques: Algorithm Development. *JMIR Med Inform* . 2020;8(6):e17364. doi: 10.2196/17364. [PMC Free Article] [PMC free article] [PubMed] [CrossRef] [Google Scholar]
17. Maas P, Barrdahl M, Joshi AD, Auer PL, et al. Breast cancer risk from modifiable and nonmodifiable risk factors among white women in the United States. *JAMA Oncology* . 2016;2(10):1295–302.

18. Tan, K.C., Teoh, E.J., Yu, Q. and Goh, K.C. (2009) A Hybrid Evolutionary Algorithm for Attribute Selection in Data Mining. *Journal of Expert System with Applications*, 36, 8616-8630.
19. R. L. Siegel, K. D. Miller and A. Jemal, "Cancer statistics", *Ca-a Cancer Journal for Clinicians*, vol. 68, no. 1, pp. 7-30, 2018.
20. Yao X, Rushlow DR, Inselman JW, McCoy RG, Thacher TD, Behnken EM, Bernard ME, Rosas SL, Akfaly A, Misra A, et al. Artificial intelligence-enabled electrocardiograms for identification of patients with low ejection fraction: a pragmatic, randomized clinical trial. *Nat Med*. 2021; 27:815–819.
21. Singh, Y., Bhatia, P.K., and Sangwan, O. (2007) A Review of Studies on Machine Learning Techniques. *International Journal of Computer Science and Security*, 1, 70-84.
22. Jamthikar A, Gupta D, Saba L, Khanna NN, Araki T, Viskovic K, Mavrogeni S, Laird JR, Pareek G, Miner M, et al. Cardiovascular/stroke risk predictive calculators: a comparison between statistical and machine learning models. *Cardiovasc Diagn Ther*. 2020; 10:919–938.
23. Gulia, A., Vohra, R. and Rani, P. (2014) Liver Patient Classification Using Intelligent Techniques. (IJCSIT) *International Journal of Computer Science and Information Technologies*, 5, 5110-5115.
24. Christodoulou E, Ma J, Collins GS, Steyerberg EW, Verbakel JY, Van Calster B. A systematic review shows no performance benefit of machine learning over logistic regression for clinical prediction models. *J Clin Epidemiol*. 2019; 110:12–22.
25. Brown KL, Ridout DA, Hoskote A, Verhulst L, Ricci M, Bull C. Delayed diagnosis of congenital heart disease worsens preoperative condition and outcome of surgery in neonates. *Heart*. 2006; 92:1298–1302.
26. Grogan M, Lopez-Jimenez F, Cohen-Shelly M, Dispenzieri A, Attia ZI, Abou Ezzedine OF, Lin G, Kapa S, Borgeson DD, Friedman PA, et al. Artificial intelligence-enhanced electrocardiogram for the early detection of cardiac amyloidosis. *Mayo Clin Proc*. 2021; 96:2768–2778.
27. P. Mishra, V. Varadharajan, U. Tupakula, and E. S. Pilli, "A detailed investigation and analysis of using machine learning techniques for intrusion detection," *IEEE Communications Surveys & Tutorials*, vol. 21, no. 1, pp. 686–728, 2019.
28. Z. Ma, H. Ge, Y. Liu, M. Zhao, and J. Ma, "A combination method for android malware detection based on control flow graphs and machine learning algorithms," *IEEE Access*, vol. 7, pp. 21235–21245, 2019.
29. Y. Zhang and M. Q. An, "Deep learning-and transfer learning-based super resolution reconstruction from single medical image," *Journal of Healthcare Engineering*, vol. 2017, 2017.
30. S. Ghosh, P. Shivakumara, P. Roy, U. Pal, and T. Lu, "Graphology based handwritten character analysis for human behaviour identification," *CAAI Transactions on Intelligence Technology*, vol. 5, no. 1, pp. 55–65, 2020.
31. Z. Allam, G. Dey, and D. Jones, "Artificial Intelligence (AI) provided early detection of the coronavirus (COVID-19) in China and will influence future urban health policy internationally," *AI*, vol. 1, no. 2, pp. 156–165, 2020.