Original Article

Prediction Of The Sickle Cell Anaemia Disease Using Machine Learning Techniques

Shivangi Shekhar

Kalinga University, Bhubaneswar, India
Email: [1] shivangishekhar5@gmail.com
DOI: 10.47750/pnr.2022.13.S10.371

Abstract

This research examines the utilization of machine learning to classify medical datasets, especially to guide sickle cell illness therapy. Numerous studies had shown that machine learning algorithms enhance pre-processing of medical time-series data signals and help classify medical data accurately. This study presents data for different kinds of medical learning algorithms. The first case is to identify drug dosages for individuals with Sickle Cell Disorder. The present study explores the performance and accuracy of Fuzzy C- means architectures. The major goal of using categorization is to help healthcare institutions give proper medicine dosage. Accuracy curves for the training and testing datasets are represented by the matching curves for each of the bars on the graphs During trials, the Fuzzy C-means delivered the best overall results with an accuracy of 99.90%.

Keywords: Machine Learning, Sickle Cell Anaemia, Fuzzy C-Means, Healthcare, Drug Dosages, Medical Dataset

I. INTRODUCTION

In the natural world, blood circulates in a small circular pattern, delivering oxygen to structures since spherical portions of the person's body. Natural blood has a lifetime of around 120 days, and each blood cell replaces the previous one every 120 days [1]. Generally, civilization faces a critical shortage of healthcare resources, both human and technological. The disaster was very well in the outbreak, especially in rural areas where there aren't enough resources to administer healthcare. Due to a lack of doctors in rural areas, many rural residents have died as a result of clinical errors made by the doctors [2].

Anemia is a kind of blood syndrome in which the native hemoglobin in red blood cells changes structurally, resulting in red blood cell enlargement. Sickle cell disease manifests as a sticky and hard Disc shape, resulting in the abstinence of blood supply to the human body [3]. A genetic RBC disorder characterized by a deficit in hemoglobin S is also known as sickle cell anemia [4]. When HbS components are present inside RBCs become bigger because there isn't enough oxygen, It significantly alters RBC form, elasticity, and adhesion characteristics. Furthermore, RBCs stiffen and become more robust. with a wide range of influences in the cell population [5]. Recent advances in medical imaging and computational image processing techniques may aid in the monitoring of SCD patients' health. This is what Darrow et al., as well as others, believe [6]. Sickle cell disease is a damaging hereditary disease that impacts a large number of individuals. This condition impacts the health and living standards of people with sickle cell disease because of unusual red blood cells (RBCs). Although SCD seems to be relatively difficult, there are many treatment outcomes that range from early death to illness that is almost completely undetectable by the eye.

The major cause of the condition in those who have it is a collection of inherited disorders that have modified a protein called hemoglobin within the RBC. There are 7 million babies born each year with a congenital defect or a disease that is passed on from generation from their families, the WHO says [7]. The optimization algorithm for the heart is a one-of-a-kind optimization algorithm. In order to create the heart algorithm, the human heart and respiratory system were studied. This method starts with a random manner of candidate solutions and then computes the optimal solution for each of those candidates' solutions. As a result, one solution is designated as a heart, while the others are designated as blood [8]. According to a report from the National Health Service, there are 250,000 people in the United Kingdom who do have sickle cell disease [9]. Sickle cell Blood flow is shown in the figure 1.

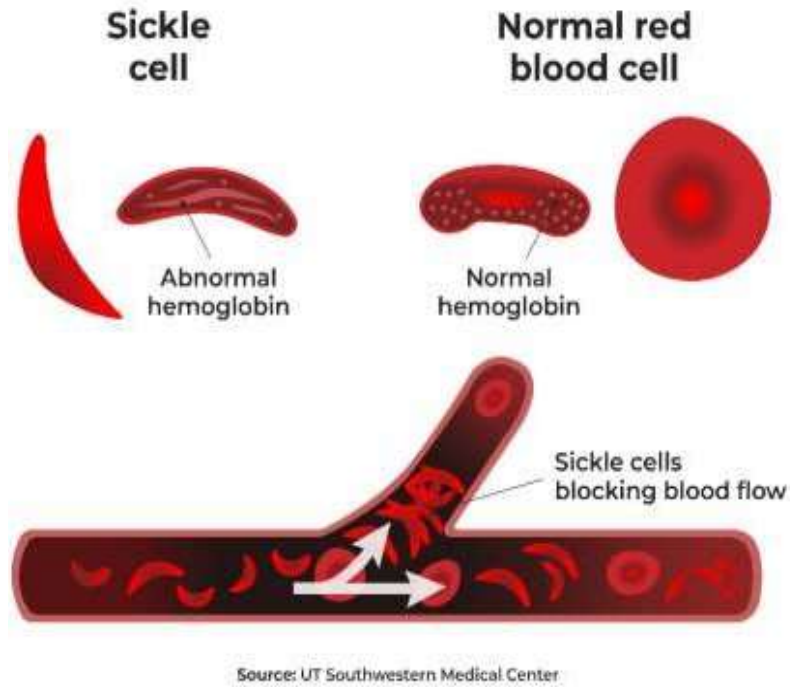

Figure 1: Blood flow of sickle cells [10]

1.1. Symptoms of sickle cell anemia

Symptoms of sickle cell Anemia occur when there are not enough red blood cells in your body, as a result of the following sickle cell illness is indicated by a difference in the morphology of the red blood cell in the patient. Manually inspecting microscopic photographs is a time-exhausting and difficult method.

- The first signs and symptoms of sickle cell anemia appear around the age of five months. Individuals vary and change with time, and this is valid for them as well. Symptoms and indications might include a variety of things.
- The RBC prevents sickle cell blood from traveling via tiny blood channels and into pulmonary capillaries. When such cells do not give oxygen to the eyes, it is possible that macular damage may occur. It has a negative impact on one's foresight [11].

1.2 Sickle Cell Impediments

Cardiac stroke may occur when sickle cells Impediments cause a disruption in the blood flow to the brain, resulting in redness and swelling. The symptoms of a stroke include weariness or numbness in the arms and legs, frequent speech difficulty, and loss of vision.

- Individuals with sickle cell anemia may develop pulmonary (high blood pressure in the lungs). Neither children nor adults are immune to the illness's effects. Breathing difficulties and fatigue are possible complications of the condition.
- SCA (sickle cell anemia) may create open sores, which can contribute to the development of ulcers on the legs [11].

1.3 Formulation of Sickle cell

The genetic factor that controls the production of hemoglobin, a steam-powerful substance that gives blood its red color and allows red blood cells to move oxygen from the lungs to the rest of the body, is mutated in this disorder, causing it (hemoglobin) [11].

In the figure 2, you can see the normal red blood cells and the sickle red blood cells together.

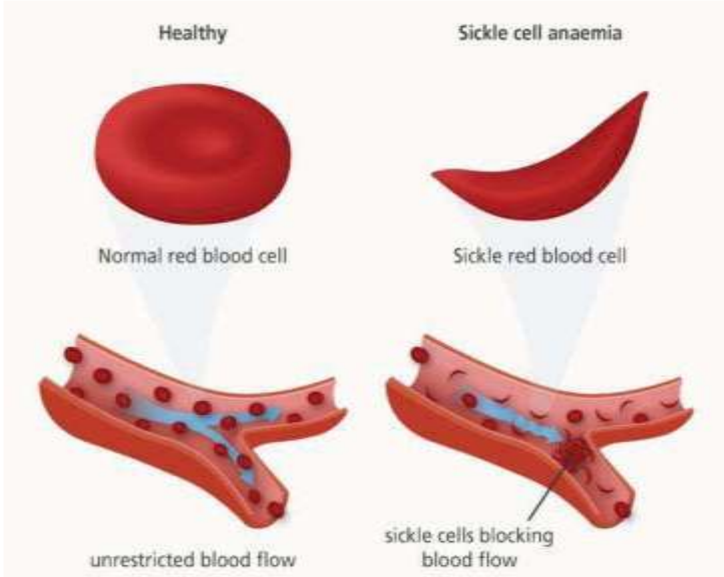

Figure *2*: The Normal Red Blood Cells and Sickle Red Blood Cells [12]

1.4 Recognition of sickle cell

Sickle cell disease is a difficult disease to detect and classify accurately in automated medical diagnosis. To analyze several techniques for improvement/renewal, segmentation, and categorization that are used to detect sickle cell disease, as mentioned below [13].

- SVM-based characterization of Sickle Cell Disease may liver failure is an irreversible disease that often leads to hepatocellular carcinoma. Early diagnosis aids neurologists in disease management and surgical planning. The present work uses an SVM classifier to characterize normal and cirrhotic liver [12].
- Sickle cell disease identification is vital in Medical Image Analysis. Utilize advanced sickle systems can be found, plan treatment evaluation results treatment. The shape of a normal blood cell versus a sickle cell differs markedly [14]. Prediction of "sudden cardiac death (SCD)". Future research should focus on predicting the development of SCD using heart rate variability (HRV) signals. All clinically significant features are then ranked and fed into classifiers such as K-nearest neighbor (KNN), decision tree (DT), and support vector machine (SVM) [15].
- A significant source of mortality in patients with SCD has been identified in various parts of the world, with the burden being particularly high in Sub- Saharan Africa. It may use a multi-algorithm, the composite may be used to predict the risk of anemia in pediatric SCD affected roles. Using demographical data, the research discover that could predict whether or not a patient with SCD will suffer from anemia [16].

II. LITERATURE REVIEW

This part is reviewed on Sickle Cell Disease. There is a wide range of authors who used the different techniques and presented their findings, as mentioned below.

Ran An et al, 2020[17] explained the reason for the article, One-third of the global total suffers from anemia, with women and children suffering the brunt of the condition. According to the World Health Organization (WHO), anemia may result in developmental delays in kids, as well as a high prevalence of disease and humanity in those who have it. Sickle cell disease and other inherited hemoglobin (Hb) abnormalities are two examples of this kind of thing that can happen to people with sickle cell disease. As a consequence, it is critical to ensure that blood Hb levels, anemia status, and Hb variations are all examined and monitored concurrently. This is particularly critical in places where anemia and genetic hemoglobin abnormalities are prevalent. Hb electrophoresis is a critical in vitro diagnostic (IVD) technique for evaluating whether an individual has sickle cell illness or the sickle cell trait. Additionally, it covers the recognition and observation of blood hemoglobin levels, anemia status, and the various forms of hemoglobin in the blood. Computer vision and deep learning techniques may be utilized to obtain novel data from Hb electrophoresis.

Nataˇsa Petrovic et al, 2020[18] explained that it has been suggested that for each classifier, believe the behavior of the proposed system that uses Randomized and Grid search should be used as a guide. Online, as made the limitations for each classifier, the code archive that was used, the uncertainty substances that were used with the natural data, and a database of erythrocytes that could be used to test the results available. For future study, There are a lot of different diseases that can be diagnosed with cell morphology analysis from microscope images, and also want to see if this can help to figure them out. Different classifiers were said to be good at a certain task. Support vector machines, selection trees, regression trees, extra trees, and gradient boosting are just a few of the tasks that it is capable of carrying out. Additionally, it is capable of doing tasks such as detecting the k-nearest neighbors and calculating the route between two multilayer perceptron (MLP). Among the areas of research we are considering are ensemble voting, feature subsets for improved classifying, feature selection for sickle-cell disease diagnostic support, and interpretability based on PCA and LDA, as well as how to evaluate the results of the research undertaken.

Laith Alzubaidi et al, 2020[19] suggested that it is the primary function of red blood cells (RBCs) to deliver oxygen to all regions of the body throughout the day. When someone develops sickle cell anemia, their red blood cells (RBCs) take on the appearance of a crescent or sickle. Sickle cell syndrome is affected by platelets in the blood that aggregates in blood tubes and cause them to get clogged. This results in discomfort and, in extraordinary situations, death. In some locations, it is known as "abnormal hemoglobin". An artificial neural network structure called a Deep Convolutional Neural System (DCNN) was designed to organize red blood cells into three categories: normal, abnormal (sickle cell anemia type), and other. A range of classifiers, including the Support Vector Machine (SVM), the K- Nearest Neighbor (KNN), and the Extreme Learning Machine (ELM), have been shown to perform exceptionally well in particular contexts. Convolutional Neural Networks (CNNs) have recently been used to categorize red blood cells (RBC) and determine whether or not anyone has sickle cell disease. In future work, the post model will be utilized to classify white blood cell diseases.

Hajara Abdelkarim Aliyu1 et al. 2019[20] explained Sickle cell anemia (SCA) is a dangerous bloodline infection that can cause people to be hospitalized a lot and even kill them. People with SCA have to stay in the hospital a lot and may die. When looking at SCA patient blood videos under a microscope, the physical process of finding and sorting abnormal cells is time-consuming, laborious, and easy to make mistakes. It needs the expertise of a qualified hematologist to do well, and it takes a lot of practice. We used a tool called the Red Blood Cell (RBC) classifier to do well at a certain job. As a result, The specificity is decreased because there are fewer normal cells in the blood spread pictures of SCA patients.

Fan Yang et al, 2018[21] suggested Controlling pain is crucial in treating Sickle Cell Disease. It is critical to do an appropriate pain assessment. The beginning stage of pain management is the most critical. On the other hand, pain is a subjective sensation that is difficult to quantify objectively. The inter-individual accuracy has been increased to 0.681 percent. As a result, a straightforward strategy such as KNN is a little less able to impress alternative strategies on this dataset. Unsurprisingly, KNN performed the lowest, given that it only analyzed the immediate area of the current data point. Predictive models will be used in the future to increase the system's overall performance by including patient-specific and demographic data such as age, gender, and baseline pain score. The use of pain medications will be taken into consideration during the pain change. There are many categorization algorithms available, including Naive Bayes (NB), The SVM, the Multinomial Regression (MLR), the K-Nearest Neighbors (KNN), and the Random Forests (RF) were secondhand.

Laith Alzubaidi et al, 2018[22] explained that Sickle Cell Anemia, damaged red blood cells causes blood vessels to become clogged, resulting in the formation of scar tissue. Results in discomfort and, in extraordinary situations, death. In certain circles, it's referred to as "abnormal hemoglobin". As formulated a system for categorizing red blood cells into three categories: normal, abnormal (sickle cell anemia variety), and other (unknown). A deep Convolution Network architecture was employed. A classifier named ECOC was placed on top of our model to increase the effectiveness of our model's test patches when used in combination with the ECOC classifier. Red blood cells in sickle cell Affected roles are classified according to their size. Many other sorts of classification techniques were utilized, including the SVM Classifier, Convolutional Neural Network (CNNs), and other forms of classification methods Aside from that, the manual approach is too tough to utilize since the edge, position, form, and size may all vary significantly. It will also provide you with a solid indication of your risk of developing sickle cell anemia. Sickle cell anemia may be treated with the purpose of image processing and machine learning methods.

Mohammed Khalaf et al, 2017[23] suggested a lot of information about how to organize medical data, especially for Sickle Cell disease-modifying therapies. It lays out how many treatments each person with a certain disease needs to be given to help them get better. It is a form of blood illness that may impair the body's capacity to fight against infections and can be fatal if untreated. The main goal of categorization is to make healthcare more accessible. People and groups work together to make sure that the right amount of medicine is given. A lot of different techniques to classify things, such as the Jordan Neural Network (JNN), Elman-Jordan Hybrid Neural Network (EJNN), Elman Neural Network (ENN), Random Forest, Decision Tree

Concert Classifier, Random Lookup Model, SVM, Linear Combiner Network among others, are examples of machine learning techniques. During our experiments, as looked at a lot of different models and saw that a few of them did work well.

Dhafar Hamed Abd et al. 2017[11] explained about patients with Sickle Cell Disease (SCD) need to be tested on a routine basis. Real-time healthcare monitoring is becoming more and more important because people want to get help from home and get it from people who know them. Use Sequential Minimal Optimization and Rules Java Repeated Incremental Pruning (JRip) with Support Vector Machine (SVM) to do the classification. Stump tree decision making, Naive Bays, and Naive Bays for Comparative analysis are all methods that are used. As a consequence, persons who have Sickle Cell Disease (SCD) in people who haven't been afflicted are more likely to get the disease. It was determined that SMO was the most accurate and had the lowest mistake rate when it came to classification.

Mengjia Xu et al, 2017[24] explained about the hematological condition called sickle cell disease (SCD) causes blood vessels to become blocked, which can cause pain and even death. It's also called sickle cell disease. Human red blood cells (RBCs) from persons with SCD exhibit a wide range of shapes and sizes, as well as significant biomechanical and bio-rheological properties. Their intensity, vulnerability, and adhesive properties, among other things, are all significant aspects to take into consideration. As a consequence, having an actual and effective method of assessing the form and classification is an advantage of RBCs can help people who have cancer learn more about their condition and improve their chances of having the disease go away. dCNNs and Red Blood Cell (RBC) classifiers were used to do well at a certain project. Furthermore, can look into how to quantify the specific shape on the categorized RBC image data and to do a broad multiscale structure analysis on the set of data.

2.1 Comparison of the reviewed literature

There is a wide range of authors who used the techniques and presented their discoveries, as can be seen in table 1.

Table 1. Summary of literature review

S.No.	Author	Technology Used	Outcomes	Research Gap
1	Ran An et al., 2020 [17]	IVD	As a result, Computer vision and machine learning techniques appear to be effective at extracting novel information from hemoglobin electrophoresis.	
2	Nataˇsa Petrovic et al., 2020 [18]	SVM, DT, RF, ET, GB, KNN, MLP	As a result, research are designed to verify the most efficient classifier for helping in the diagnosis of SCD.	Can cell morphology analysis from microscope pictures be used to help doctors figure out what kinds of illnesses they have.
3	Laith Alzubaidi et al., 2020 [19]	SVM, KNN, ELM, CNNs	As a result, A clear picture has been created of the severity of sickle cell anemia's threat level.	When it comes time to identify white blood disorders, we want to employ our previously trained model.
4	Hajara Abdelkarim Aliyu1, 2019 [20]	RBC	There aren't as many healthy cells in the blood smudge pictures of SCA patients, the specificity isn't as high as it should be.	
5	Fan Yang et al., 2018 [21]	MLR, NB, SVM, KNN, RF	As a result, A prototype machine learning standard for guessing pain levels in SCD patients has been developed.	In the future, information such as age, gender, and a control pain score will be added.

6	Laith Alzubaidi et al., 2018 [22]	SVM, CNNs	For the classification job, The error-correcting output codes (ECOC) classification is produced as a result of this process.	It is vital to accurately categories RBCs in order to get an accurate cell count. A clear picture of the amount of risk associated with sickle cell anemia will be provided as a conclusion.
7	Mohammed Khalaf et al., 2017 [23]	ENN, RFC, SVM, LNN, ROM, JNN, EJNN, LEVNN	During our trials, the findings of a variety of models revealed that among the suggested models	The use of global optimization methods like genetic optimization to investigate the space of feasible recurrent network structures more thoroughly.
8	Dhafar Hamed Abd et al., 2017 [24]	Rules, JRiP, SVM, SMO, decision, Tree and Naïve Bays.	As a consequence, people with SCD differ from those who are not affected.	
9	Mengjia Xu et al., 2017 [25]	RBC, DCNN	The result of this is a comparison of the performances of the deep CNNs applied in the various RBC division situations.	It is feasible to examine the certain shape factor quantification for the detected RBC picture information.

III. BACKGROUND STUDY

The demand for real-time healthcare monitoring systems has productivity increased. SCD patients require continual testing, monitoring, and care. These services require an integrated healthcare system. Such healthcare devices are built possible by modern information systems and technologies. This study built a method for testing, monitoring, and tracking patients' SCD. The suggested method uses supervised machine learning to evaluate patient data and inform medical personnel. For classification, SVM uses four methods: Consecutive Minimal Optimization, Java Repeated Incremental Pruning, Tree Decision Stump (TDS), and Naive Bays (NB). Many experiments were conducted using the four machine learning algorithms to differentiate SCD patients from healthy people. Using the SMO algorithm, 99 percent of categories were accurate [13].

IV. PROBLEM FORMULATION

Sickle cell anemia, often termed SCD is a hematological disorder that causes occlusion in blood vessels, leading to hurtful episodes and even death. The key function of red blood cells (erythrocytes) is to supply all the parts of the human body with oxygen. SCD cannot be cured but can be treated to ease the pain and prevent problems caused by this disease. One of the methods for diagnosing SCD is by observation of a patient’s peripheral blood samples under a microscope and counting sickle cells, which is a tedious and time-consuming task. The proposed research work implements the machine learning approach for the analysis of microscopic blood smear images. In the proposed work regression algorithm is used for the prediction of SCD detection. In this work, FCM and clustering are also used for the data normalization in the preprocessing stage.

V. RESEARCH METHODOLOGY

As part of this research, it examined existing industry and academic literature in the identification and classification of Sickle Cell Disease, as well as determining Sickle Cell Anemia that was fed into the machine learning techniques to determine what pattern predicts the behavior of infected patients of sickle cell disease.

5.1 Technique Used

The different techniques used in the proposed methodology is summarized below:

- FCM clustering algorithm The FCM clustering method, which is an evolving procedure of tuning the parameters, calculates the Centre value of each group as well as a fuzzy partition matrix. The FCM technique may not yield meaningful fuzzy partitions if the number of clusters is incorrect. In this scenario, the FCM algorithm can use cluster validity criteria to determine optimal partitions [25].

- Regression Algorithms: Regression analysis algorithms are a subtype of data analytics that makes use of a relationship between factors (goal) and independent variables. Some of the most well-known regression models are Linear Regression, Logistic Regression, Backward Elimination, Normal generalized least regression, multilayer perceptron spline and Locally Estimated Scatterplot Smoothing. It assumes that the dependent and independent variables have a linear relationship. MLR is a technique based on the least- squares concept, with the estimated parameter expressed as a linear function. [26].

5.2 Proposed Methodology

The basic block diagram of the proposed methodology is shown in the figure 3.

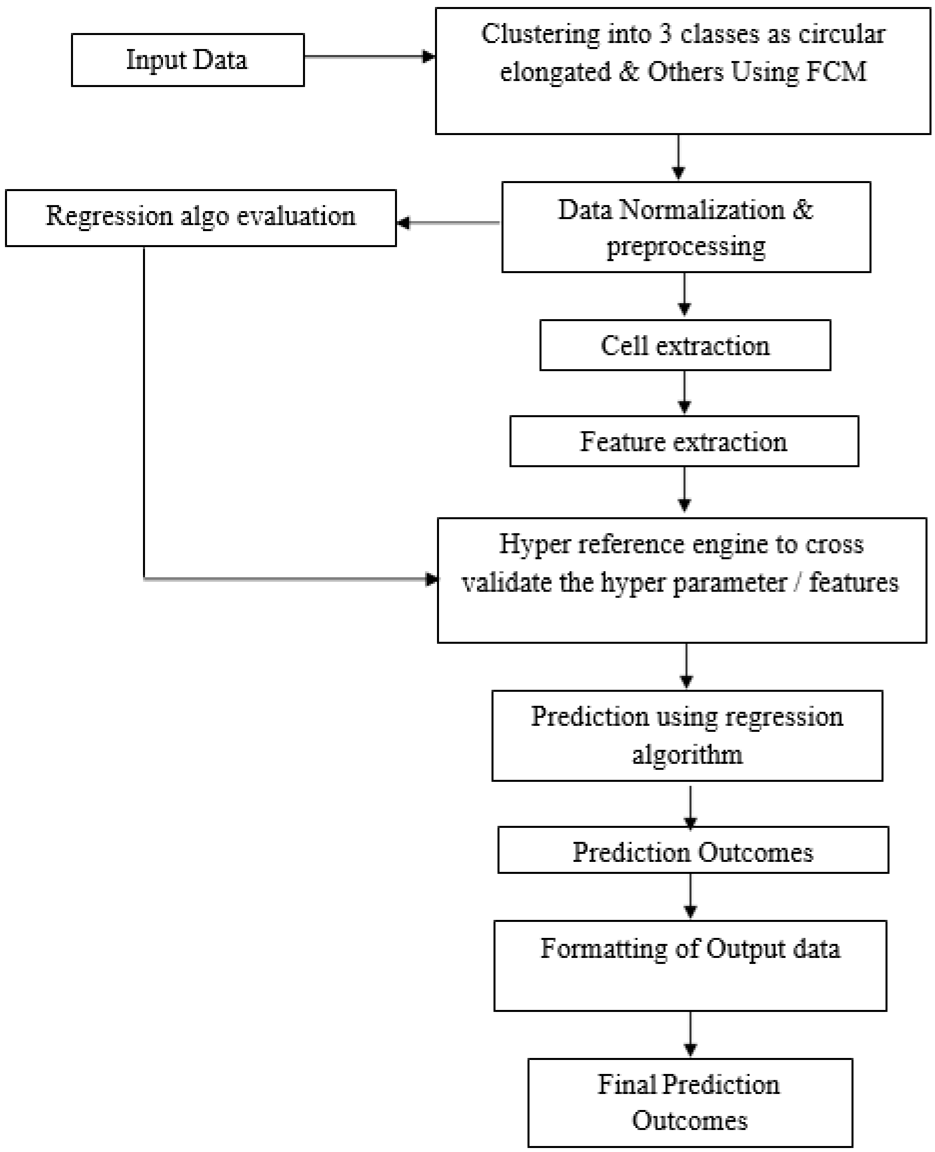

Figure 2. Block diagram of proposed methodology

The step wise explanation of the proposed methodology is shown in the following steps:

Steps 1. In the first phase, input data is used to select a location. Alder Hey Children's Hospital in Liverpool, UK, provided the dataset used for our tests on SCD patients, which was compiled during a five-year period for the purpose of research. The SCD trait is indicated by a total of 13 factors in each sample, as depicted.

Step 2. According to Clustering, the recommended sites are subjected to a second screening method in the second stage. This phase made use of the Fuzzy C-means (FCM) system to simplify phase 1 and forecast the following phase.

Step 3. The third phase it is regarded to be the most crucial since it is during this phase that the analysis and rating. This is Data Normalization & preprocessing.

Step 4. The fourth phase, it is regarded to be cell extraction. Cells were retrieved using standard image processing techniques. To begin Regression analysis is a subset of predictive analytics that takes advantage of the relationship between the dependent (target) and independent variable quantity. Some of the more well-known regression models include linear regression, logistic regression, regression analysis, ordinary least squares recurrence (OLSR), machine learning algorithms splines, and locally estimated scatterplot smoothing (LOESS). It is predicated on the existence of a linear affiliation between the dependent and independent variables. Based on the notion of least squares, MLR calculates an estimate of a parameter by using a linear function to represent it. This process begins with an image's deterioration and concludes with its dilatation. Each of these processes leads to the most essential, edge detection. The Canny filter is used to make the image's edges visible. It employs a multi-stage technique for detecting multiple edges in cell images.

Step 5. The following part of this investigation is feature extraction. The initial stage in every classification problem involving Machine Learning (ML) techniques is feature extraction. In order to differentiate the newly discovered cell kinds, as identified the most frequently used characteristics in the literature. According to their state of fine art, the retrieved features are classified into three categories. form or geometry, color, and texture.

Step 6. These indices were created and combined utilizing the Hyper reference engine in order to cross-validate the hyperparameter/characteristics. It is a significant mix of the evaluation of linear regression and the primary processes. To reduce overfitting, as the evaluated performance of the classifier using a tenfold bridge technique on both Purely random and Grid search methods [10]. The data set was divided into two components, with a 70/30 split between teaching and challenging. While the teaching set was used to fine-tune the classifiers, the testing set was used to generate predictions using the optimal combination of Randomized and Grid searches. Parameter selection also wanted to compare our classifiers' performance. It is compared each classifier performance after configuring it to its optimal parameters.

Step 7. The study's next phase is regression prediction. We used four well-known pain predictors. First, used Multinomial Logistic Regression (MLR). Predicting EHR Imputation Labels Using EHR Data to Compute (Case 1) (Case 3): Prediction of Patient Labels (Case 4) A Prediction and a B Prediction Psychoanalysis of One's Own Feelings This function is inferred to be logistic or sigmoid. Given the input predictor variables, the logistic function gives a number between 0 and 1. Logistic regression predicts a binary target variable. MLR predicts the probability of a minimum target variable belonging to more than two groups. An MLR model gives probabilities to each class to determine the most likely. One of MLR's main advantages is that it can make qualitative inferences about processes from the values and significance of each predictor variable [25]. To identify normal and abnormal erythrocytes, Parameters for the classifier weren't Only aberrant erythrocytes were used to classify SCA [22].

Step 8. In the next phase, Prediction Outcomes, First, as used MLR (MLR). EHR Imputation Pain Forecast Data from EHR Case 1: Without Patient Labels (Case 2), (Case 3), (Case 4) A-B Prediction Intrapersonal Affective Investigation It is possible to infer the exponential or quadratic function. Predictor variables are used to determine the transition probabilities of each objective variable value based on the input response variable. Logistic regression predicts a binary target variable. More than two classes for a nominal target variable are predicted by MLR, an extension of binary Logistic Regression.

Step 9. The study's final predicted outcomes are the next step. An MLR model gives probabilities to each class in order to choose the one with the highest probability. The consequences in this scenario are 11 pain notches. The primary benefits of MLR are its simplicity, speed, and capacity to derive subjective judgements about phenomena built on the values and importance of each interpreter changing.

VI. RESULTS AND DISCUSSION

The findings of the testing environment utilized in the suggested trials of the setup of every system, and ultimately the performances assessment measures employed to quantify the consequences of the machine learning models have all been described in this part. The dataset contains 1170 sample points and a unique goal parameter that describes the hydroxyurea/hydroxycarbamide prescription dose in milligrams. To aid this categorization analysis, discretization of the target dose into three bins has been done, designated classes 1 through 3, constructed by reducing the output range (in Milligrams) in membership periods of equally sized: Class 1: [300≤ Y < 532 mg], Class 2: [532≤ Y < 765 mg], Class 3: [765≤ Y ≤1000 mg]. This split was carried out, in addition, to give enough classification distribution throughout the sample data and maintain a certain degree of accuracy for the dose output. Because the data sample was restricted to one hundred samples, making a meaningful split for more than three groups was difficult, as shown in fig 3.

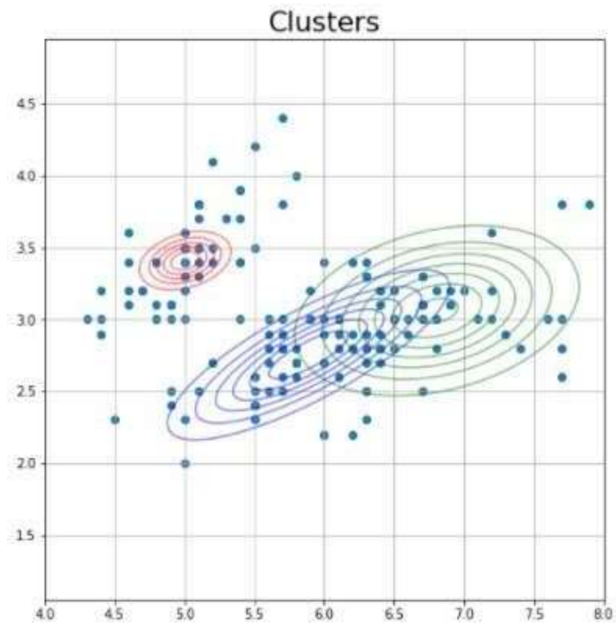

Figure 3: Classification of data in 3 different clusters.

A segment of the datasets is used for learning, a segment of the dataset is used for verification, and a portion is used for evaluating Using the holdout approach, testing whether the results of a statistical study transfer to a different dataset has been done. Dataset segmentation was employed in this research to establish a mean proportion of right classifications. It is possible to compare the accuracy curves for different models by comparing their accuracy bars.

Classifiers developed in this study beat all alternative classifiers, even RNN-based classification systems. There is a perfect match between the suggested classifiers and all of the trained data for all of the operational points. During the training of class 1, the accuracy of the model is 09569377990430622, sensitivity is 09615384615384616,

Specificity is 0.9523809523809523, precision is 0.9615384615384616, recall is 0.9523809523809523, F Score is 0.9569377990430622, J Score is 0.9139194139194138. whereas for the training of class 2 model the accuracy is 0.8888888888888888, sensitivity is 0.8928571428571429, specificity is 0.8849557522123894, precision is 0.8928571428571429, Recall is 0.889557522123894, F Score is 0.8888888888888891, J Score is

0.7778128950695322. and for class 3 model training the accuracy is 0.970873786407767, sensitivity is 0.970873786407767, specificity is 0. 970873786407767, precision is 0. 970873786407767, Recall is 0.

970873786407767, F Score is 0. 970873786407767, J Score is 0.0.941747572815534 as shown in Fig 4.

```
Training results
For Class 1

Accuracy is  :  0.9569377990430622
Sensitivity is  :  0.9615384615384616
Specificity is :  0.9523809523809523
Precision is :  0.9615384615384616
Recall is :  0.9523809523809523
fscore is :  0.9569377990430622
jscore is :  0.9139194139194138

For Class 2

Accuracy is  :  0.8888888888888888
Sensitivity is  :  0.8928571428571429
Specificity is :  0.8849557522123894
Precision is :  0.8928571428571429
Recall is :  0.8849557522123894
fscore is :  0.8888888888888891
jscore is :  0.7778128950695322

For Class 3

Accuracy is  :  0.970873786407767
Sensitivity is  :  0.970873786407767
Specificity is :  0.970873786407767
Precision is :  0.970873786407767
Recall is :  0.970873786407767
fscore is :  0.970873786407767
jscore is :  0.941747572815534
```

Fig 4: Training results for class 1, class 2, and class3.

The results obtained for the testing of class 1, the accuracy of the model is 0.9302325581395349, sensitivity is 0.9090909090909091, Specificity is 0.9523809523809523, precision is 0.0.9090909090909091, recall is 0.0.9523809523809523, F Score is 0.9302325581395349, J Score is 0.8614718614718614. whereas for the testing of the class 2 model the accuracy is 0.8928571428571429, sensitivity is 0.8928571428571429, specificity is 0. 8928571428571429, precision is 0.8928571428571429, Recall is 0.8928571428571429, F Score is

0.8928571428571429, J Score is 0.7857142857142858. and for class 3 model testing the accuracy is 0.9900990099009901, sensitivity is 0. 9900990099009901, specificity is 0.9900990099009901, precision is 0.

9900990099009901, Recall is 0. 9900990099009901, F Score is 0. 9900990099009901, J Score is

0.0.9801980198019802 as shown in Fig 5.

```
Testing results
For Class 1

Accuracy is  :  0.9302325581395349
Sensitivity is  :  0.9090909090909091
Specificity is :  0.9523809523809523
Precision is :  0.9090909090909091
Recall is :  0.9523809523809523
fscore is :  0.9302325581395349
jscore is :  0.8614718614718613

For Class 2

Accuracy is  :  0.8928571428571429
Sensitivity is  :  0.8928571428571429
Specificity is :  0.8928571428571429
Precision is :  0.8928571428571429
Recall is :  0.8928571428571429
fscore is :  0.8928571428571429
jscore is :  0.7857142857142858

For Class 3

Accuracy is  :  0.9900990099009901
Sensitivity is  :  0.9900990099009901
Specificity is :  0.9900990099009901
Precision is :  0.9900990099009901
Recall is :  0.9900990099009901
fscore is :  0.9900990099009901
jscore is :  0.9801980198019802
```

Fig 5: Testing results for class 1, class 2 and class3

Furthermore, accuracy ranged from 0.998 suggesting strong generalization to the testing sample for the suggested classifiers' training results. Indicating a high upper limit on identification ability, such two classifiers exhibit a robust generalization of the given data information richness sources. SVM detector trials reveal that the suggested class of model is much less able of categorizing datasets than the SVM classifier itself. The area under the accuracy for every class is shown in Figs. 6 and 7 for every classifier. According to Fig. 6 and Fig. 7, the training data set and the test set are shown, respectively.

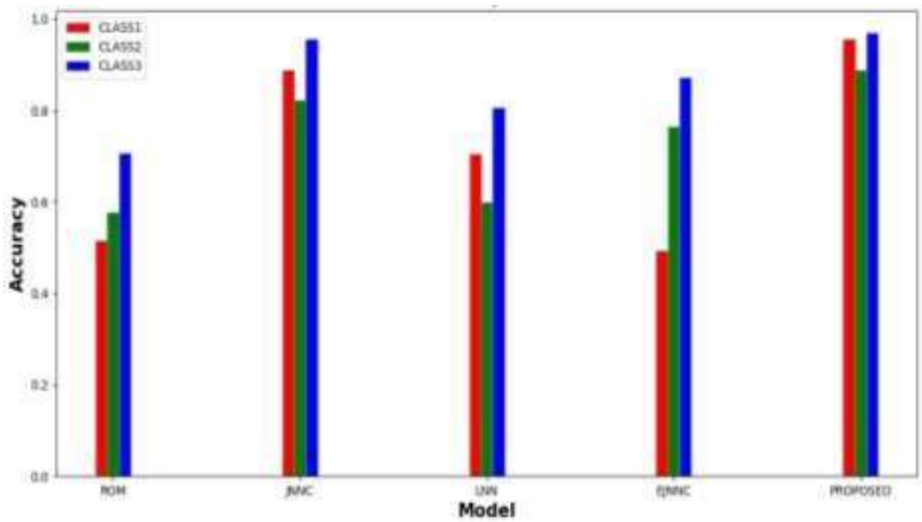

Fig 6: Trained accuracy per model

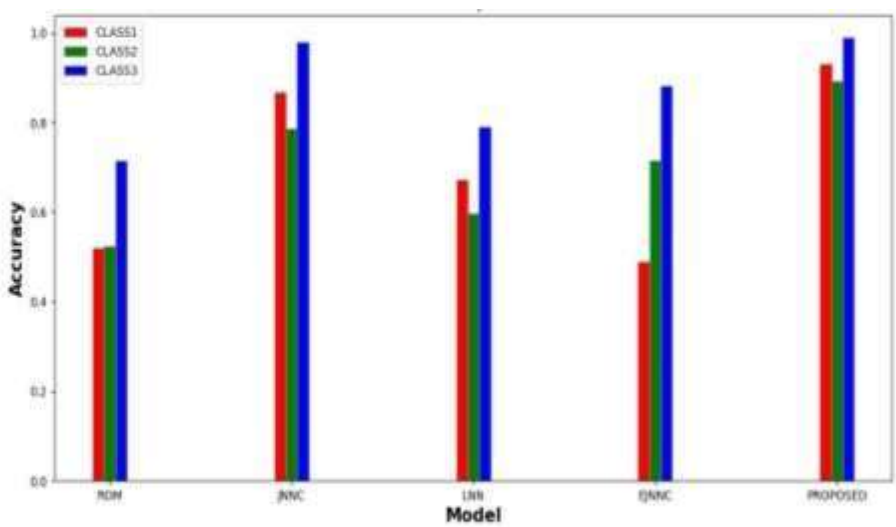

Fig 7: Test accuracy per model

It is possible to describe the worldwide competence of classification during investigation using the accuracy metric. There are two plots, one for Fig 6 and one for Fig 7, in which a model's accuracy is shown on the X-axis and the class's accuracy on the Y. An accuracy of 1 indicates a perfect classifier, whereas the accuracy of 0.5 reflects a randomized performance of the classifier. The Accuracy curve for the training and testing datasets is represented by the matching curves for each of the bars on the graphs shown below. The plot's goal is to display the accuracy values graphically so that a comparison can be made. In general, the findings show that medical data may be used to classify SCD dose levels. The selection of framework is important in attaining an acceptable outcome, as seen by the differences in performances amongst the frameworks utilized in the proposed investigation. Using the SCD data, the suggested classifier was able to correctly classify SCD patients.

VII. CONCLUSION AND FUTURE SCORE

For the categorization of SCD effective dose levels, an empirical analysis was conducted using several machine learning algorithms. In contrast to standard pharmaceutical approaches, the suggested investigation has incorporated several forms of recurrent neural networks for the analysis of medical time series gathered from SCD patients. For

preprocessing of medical time series data outputs before the categorization of medical data, existing research has proven that machine learning techniques display great efficacy. SCD effective dose levels might be classified using a machine learning technique that used artificial neural networks, according to the suggested research. Despite RNNs being able of giving some level of fitting and generalization, empirical research using patient sample data and comparator systems such as SVM and RFC indicated that RNNs are unsatisfactory in the classification context.

In the future, to investigate a wider range of potential recurrent networks Genetic optimization methods might be used. A machine learning model approach may be employed to broaden the breadth and size of this research since the present work only addresses a restricted range of designs, that may not disclose the maximum capabilities of RNNs in the assessment situation.

REFERENCES

[1] Westerman, Maxwell, and John B. Porter. "Red blood cell-derived microparticles: an overview." Blood Cells, Molecules, and Diseases 59 (2016): 134-139.

[2] Das, Sumit, and Manas K. Sanyal. "Application of AI and soft computing in healthcare: a review and speculation." vol 8 (2020): 21.

[3] Pain, Diagnosing Acute SCD, and Treating Acute SCD Pain. "Sickle Cell Disease: Core Concepts for the Emergency Physician".

[4] Anglin, Carlita. "Sickle cell disease." Journal of Consumer Health on the Internet 19, no. 2 (2015): 122-131.

[5] Fasano, Ross M., Garrett S. Booth, Megan Miles, Liping Du, Tatsuki Koyama, Emily Riehm Meier, and Naomi LC Luban. "Red blood cell alloimmunization is influenced by recipient inflammatory state at time of transfusion in patients with sickle cell disease." British journal of haematology 168, no. 2 (2015): 291-300.

[6] Darrow, Michele C., Yujin Zhang, Bertrand P. Cinquin, Elizabeth A. Smith, Rosanne Boudreau, Ryan H. Rochat, Michael F. Schmid, Yang Xia, Carolyn A. Larabell, and Wah Chiu. "Visualizing red blood cell sickling and the effects of inhibition of sphingosine kinase 1 using soft X-ray tomography." Journal of cell science 129, no. 18 (2016): 3511-3517.

[7] Sebastiani, Paola, Marco F. Ramoni, Vikki Nolan, Clinton T. Baldwin, and Martin H. Steinberg. "Genetic dissection and prognostic modeling of overt stroke in sickle cell anemia." Nature genetics 37, no. 4 (2005): 435-440.

[8] Hatamlou, Abdolreza, and Elham Ghaniyarlou. "Solving knapsack problems using heart algorithm." International Journal of Artificial Intelligence and Soft Computing 5, no. 4 (2016): 285-293.

[9] Eleftheriou, Androulla, Michael Angastiniotis, Demitrios Loukopoulos, Christos Kattamis, and John Meletis. "3rd Pan-European Conference on Haemoglobinopathies and Rare Anaemias, 24–26 October 2012, Limassol, Cyprus." Thalassemia Reports 2, no. s2 (2012): 1-47.

[10] Yeruva, Sagar, M. Sharada Varalakshmi, B. Pavan Gowtham, Y. Hari Chandana, and PESN Krishna Prasad. "Identification of sickle cell anemia using deep neural networks." Emerging Science Journal 5, no. 2 (2021): 200- 210.

[11] Abd, Dhafar Hamed, and Intisar Shadeed Al-Mejibli. "Monitoring system for sickle cell disease patients by using supervised machine learning." In 2017 Second Al-Sadiq International Conference on Multidisciplinary in IT and Communication Science and Applications (AIC-MITCSA), pp. 119-124. IEEE, 2017.

[12] Das, Pradeep Kumar, Sukadev Meher, Rutuparna Panda, and Ajith Abraham. "A review of automated methods for the detection of sickle cell disease." IEEE reviews in biomedical engineering 13 (2019): 309-324.

[13] Virmani, Jitendra, Vinod Kumar, Naveen Kalra, and Niranjan Khandelwal. "SVM-based characterisation of liver cirrhosis by singular value decomposition of GLCM matrix." International Journal of Artificial Intelligence and Soft Computing 3, no. 3 (2013): 276-296.

[14] Fujita, Hamido, U. Rajendra Acharya, Vidya K. Sudarshan, Dhanjoo N. Ghista, S. Vinitha Sree, Lim Wei Jie Eugene, and Joel EW Koh. "Sudden cardiac death (SCD) prediction based on nonlinear heart rate variability features and SCD index." Applied Soft Computing 43 (2016): 510-519.

[15] Balogun, Jeremiah Ademola, Adanze O. Asinobi, Olawale Olaniyi, Samuel Ademola Adegoke, Florence Alaba Oladeji, and Peter Adebayo Idowu. "Ensemble Model for the Risk of Anemia in Pediatric Patients With Sickle Cell Disorder." International Journal of Computers in Clinical Practice (IJCCP) 4, no. 2 (2019): 33-59.

[16] An, Ran, Yuncheng Man, Shamreen Iram, Erdem Kucukal, Muhammad Noman Hasan, Ambar Solis-Fuentes, Allison Bode, et al. "Computer Vision and Deep Learning Assisted Microchip Electrophoresis for Integrated Anemia and Sickle Cell Disease Screening." Blood 136 (2020): 46-47.

[17] Petrović, Nataša, Gabriel Moyà-Alcover, Antoni Jaume-i-Capó, and Manuel González-Hidalgo. "Sickle-cell disease diagnosis support selecting the most appropriate machine learning method: Towards a general and interpretable approach for cell morphology analysis from microscopy images." Computers in Biology and Medicine 126 (2020): 104027.

[18] Alzubaidi, Laith, Omran Al-Shamma, Mohammed A. Fadhel, Laith Farhan, and Jinglan Zhang. "Classification of red blood cells in sickle cell anemia using deep convolutional neural network." In International Conference on Intelligent Systems Design and Applications, pp. 550-559. Springer, Cham, 2018.

[19] Aliyu, Hajara Abdelkarim, Mohd Azhar Abdul Razak, Rubita Sudirman, and Norhafizah Ramli. "A deep learning AlexNet model for classification of red blood cells in sickle cell anemia." Int J Artif Intell 9, no. 2 (2020): 221-228.

[20] Yang, Fan, Tanvi Banerjee, Kalindi Narine, and Nirmish Shah. "Improving pain management in patients with sickle cell disease from physiological measures using machine learning techniques." Smart Health 7 (2018): 48-59.

[21] Khalaf, Mohammed, Abir Jaafar Hussain, Robert Keight, Dhiya Al-Jumeily, Paul Fergus, Russell Keenan, and Posco Tso. "Machine learning approaches to the application of disease-modifying therapy for a sickle cell using classification models." Neurocomputing 228 (2017): 154-164.

[22] Xu, Mengjia, Dimitrios P. Papageorgiou, Sabia Z. Abidi, Ming Dao, Hong Zhao, and George Em Karniadakis. "A deep convolutional neural network for classification of red blood cells in sickle cell anemia." PLoS computational biology 13, no. 10 (2017): e1005746.

[23] Anand, S. Krishna, TG Sundara Raman, and S. Subramanian. "Implementing a neuro-fuzzy expert system for optimising the performance of chemical recovery boiler." International Journal of Artificial Intelligence and Soft Computing 4, no. 2-3 (2014): 249-263.

[24] Gaya, Muhammad Sani, Sani Isah Abba, Muhammad Abdu Aliyu, Abubakar Ibrahim Tukur, Mubarak Auwal Saleh, Parvaneh Esmaili, and Norhaliza Abdul Wahab. "Estimation of water quality index using artificial intelligence approaches and multi-linear regression." IAES International Journal of Artificial Intelligence 9, no. 1 (2020): 126.