EVALUATION OF CHRONIC MYELOGENOUS LEUKEMIA (CML) AS THE CHRONIC PHASE OF DISEASE USING MACHINE LEARNING TECHNIQUES

Mithilesh Kumar Pandey

Research Scholar

Department of Computer Applications Veer Bahadur Singh Purvanchal University, Jaunpur

Saurabh Pal

Associate Professor

Department of Computer Applications Veer Bahadur Singh Purvanchal University, Jaunpur

Abstract:

Background*-Leukemia is a heterogeneous category of hematologic malignancies that create an increased quantity of aberrant or immature white blood cells by bone marrow or other blood-forming organs and suppress the growth of normal cells. Chronic myeloid leukaemia (CML) is an acquired myeloprulferative illness that affects the hematopoietic stem cell. Leukemia is mostly used in machine learning methods, whether it's for classifying various leukaemias or for detecting whether leukaemia is present in a patient.*

Methods*: In this paper we use Vector Machines, Nearest Neighbor, Naïve Bayes and Deep Learning algorithms for evaluating chronic myelogic leukaemiaas the chronic phase of disease.The SFS approach is used to compare data set accuracy with all features and classifier accuracy with the characteristics that are selected.*

Result: *The results indicated that the maximum accuracy is RF sfs, KNN sves, SVC rbf and SVC sfs, which is 97.66 per cent. They are successful and can forecast the development of CML disease.*

Conclusion: *For the classification of benign and CML disease, some information mining approaches are used.*

Keywords: Leukemia, Chronic, myeloid, Machine Learning, blood, etc.

1. INTRODUCTION

The term "chronic" signifies a slower progression of cancer than acute forms of leukaemia in chronic myelogen leukaemia. The term "myelogenous" in chronic myelogenous leukaemia refers to the cell type affected. The term "myelogenous" refers to the cell type of the cell. Chronic myelogenous leukaemia (CML) is unusual types of bone marrow cancer – a spongy tissue within the bones in which blood cells are produced. The number of white blood cell in the blood increased because of CML. Chronic myelogenous leukaemia may also be referred to as chronic, granulocytic, and myeloid leukaemia. It mainly affects older persons and seldom happens in youngsters, but can take place at any age [1]. The CML is an illness that is characterised by the growth in the propagation of granulocytic cell line without the loss of its potential to distinguish.There are thus an increased numbers of granulocytes and their immature precursors, including occasional blast cells, seen in the peripheral cell profile. CML accounts about 20 percent of all adult leukemias. CML, also known as chronica myeloid leukaemia is a myeloproliferative disorder.

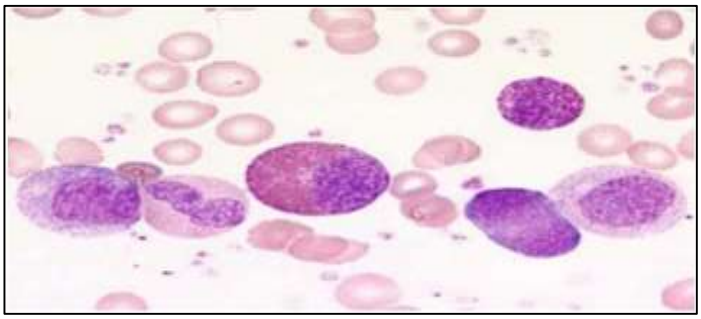

Figure 1: CML. The full granulocytic lineage containing an eosinophil and Basophil show blood film in 1000X magnification.

Chronic myeloid leukaemia is a slow-growing tissue malignancy (bone marrow). Normal bone marrow generates red blood cells (errycytes) carrying oxygen, white blood cells (leucocytes) protecting the body from infection and platelets of blood clots (thrombocytes). The bone marrow creates too many white blood cells in chronic myeloid leukaemia [2]. These cells work rather properly in the beginning. As the illness worsens, the blood and bone marrow build up to immature white blood cells called myeloblasts. The myeloblasts are overgrowing and affect the growth of other blood cells, which leads to the lack of red blood cells (anaemia) and platelets. Chronic leukaemia usually starts beyond 60 years of age. Excessive exhaustion, fever and weight loss are common characteristics. Many persons with an impaired spleen (splenomegaly) might cause the belly to feel full and lose their appetite. About half of persons with chronic leukaemia are detected when the blood test is done for another cause and have no signs and symptoms at the beginning [3]. Three phases are: the chronic phase, the accelerated phase and the explosion phase (or blast crisis). In the chronic phase, there are high numbers of mature white blood cells, and fewer than 10 percent of blood cells are myeloblasts. During this phase, the signs and symptoms are usually modest or absent and steadily worsening. The chronic phase can take months through years. In the accelerated phase, 10 to 29 percent of the blood cells comprise the number of myeloblasts somewhat higher. The symptoms and indicators continue to deteriorate. It normally takes between 4 and 6 months to accelerate, although some affected people are overrun. Blood or bone marrow cells are myeloblasts in blast crisis by 30 percent or more. Symptoms, including substantial enlargement of the spleen, bone pain as well as weight loss, are particularly severe in this phase. Uncontrolled and serious infections can pose a risk to life [4].

2. CAUSES FOR CHRONIC MYELOGENOUS LEUKEMIA (CML)

The rearrangement of genetic material (translocation) between chromosome 9 and chromosome 22 results in chronic myeloid leukaemia This translocation, which has been written as t(9;22), fuses part of the ABL1 chromosome 9 gene with part of the BCR chromosome 22 gene to create an aberrant BCR-ABL1 gene. The aberrant chromosome 22 is sometimes referred to as the chromosome Philadelphia (Philadelphia) (named for where it was first discovered) [5, 6]. The translocation is taken over the lifetime of a person and only occurs in aberrant blood cells. Somatic mutation is not inherited by this form of genetic changes. The protein function generated by the normal BCR gene is not fully known, however signalling in cells has been demonstrated to help modulate it. A number of cellular processes involving cell growth and division (proliferation), maturation (differentiation), movements (migration), and autodestruction include the protein generated by the normal gene ABL1 (apoptosis).

The aberrant protein of the fusion gene BCR-ABL1 may support cell proliferation as well as prevent apoptosis, like the ABL1 protein. Unlike ABL1, however, signals in the cell are not needed to activate it. The BCR-ABL1 protein, which is continually active, signals that the cells continue to divide themselves abnormally and inhibits the improperly producing cells, leading ultimately to the aberrant and normal blood cell shortage [7] In persons with chronic myeloid leukaemia the presence of a chromosome is a goal for molecular therapy. The BCR-ABL1 fusion genE is formed with complicated rearrangements involving other chromosomes, as well as chromosome 9 and 22 in 5 to 10 percent of the cases of chronic myeloid leukaemia. These genetic modifications are referred to as translocations of Philadelphia. This situation is comparable to that of t (9; 22). Researchers think that the progression of chronic myeloid leukaemia to accelerated phase and blast crisis plays a part in other genetic modifications. Extra copies of chromosome 8 (trisomy 8), chlomosome 17 abnormalities known as isochromosome 7 and an additional copy (double-copy) of the Philadelphia chromosome are the most common genetic alterations associated with progression to the blast crisis. In cells containing Philadelphia chromosomes, these somatic mutations will probably further encourage cell proliferation in the unregulated environment. In general, this ailment does not inherit but comes from a mutation in the cells of the body following conception. This change is known as a somatic mutation [8].

3. LITERATURE REVIEW

Cortes, J., Lang, F. (2021) [9] The BCR-ABL1 fusion protein drives chronic myeloid leukaemia (CML) and forms a translocation between chromosomes 9 and 22 that causes the chromosome of Philadelphia. The fusion protein BCR-ABL1 is an ideal target for inhibitors of tyrosine kinase (TKIs), aimed at the binding site for adenosine triphosphate (ATP) of ABL1. While the CML prognosis has been much improved, many individuals fail to receive treatment, some of which require multiple lines of CML treatment. Mutations can take place at the ATP binding point of ABL1, which causes resistance by blocking many of these medicines from binding and leaving patients with limited therapy alternatives. Adverse symptoms that can lead to withdrawal of treatment in some patients are also related with the approved TKI. Efficacy declines with each phase of therapy; results suggest minimal clinical advantage for third line (3L), inhibitor of tyrosine-kinase (2GTKI) of the second-generation treatment after failure of TKI and TKI of the first generation. For a patient population which requires therapy in 3L environments and beyond, new therapeutic alternatives are required. This examination emphasises the need for clear recommendations and innovative therapeutics for patients with and beyond 3L therapy.

Hehlmann, Rüdiger (2020) [10] New information into the optimal management of chronic myeloid leukaemia has arisen from growing industry and academia clinical trial. Unpredictably less significant than previously thought, speed of reaction does not indicate survival or is of unequivocal significance to treatment free remission (TFR). Severe and chronic toxicity with imatinib replacement inhibitors of tyrosine kinase has been seen. Generic imatinib in many patients has been cost-effectively treated first-line in chronic conditions. Previous recognition of the conclusion of genetic evaluation phase could boost odds of an blast crisis (BC). TFR has become a major new CML therapeutic objective. Recently, ELN restructured its recommendations on the treatment of CML to reflect this new reality. This paper will focus on recent breakthroughs and current evidence for CML treatment in 2020 after a short overview of 175 years of its history.

Bonifacio (2019) [11] in the era of tyrosine kinase inhibitors (TKI) treatment, management of chronic myeloid leukaemia (CML) in later stages remains an issue. Cytogenetic clonal growth and development of resistance mutations is critical in limiting the efficacy

of subsequent therapies. Accelerated (AP) or blast phase (BP) diagnoses are diagnosed in <5 per cent of the patients and therapy progressions have been drastically reduced with the accessibility of efficient treatment for chronic phase (CP). Because of the reduced number of patients, there are few randomised studies in this context and there is insufficient evidence. However, three main scenarias can be drawn: (a) the PA diagnoses are more likely than the CP patients to fail, but their outcome is similarly favourable if they are optimally treated with frontline TKI therapy; (b) the BP diagnoses can be treated either with TKI alone or with the TKI with the conventional chemotherapy schemes as well as with subsequent transplant decisions;

Thompson et al (2015) [12] Landmark work followed, recognising the underlying translocation between chromosomes 9 and 22 that gave rise to this anomaly and shortly subsequently, the individual genes involved and the pathophysiologic consequences of this unique rearrangement. Few neoplastic disorders have experienced a change in a short time like chronic myeloid leukaemia (CML) has in the recent few years. In 1960, CML was the first malignancy where a unique chromosomal aberration, "a minute chromosome", was detected and a pathophysiologic association indicated. The broad usage of tyrosine kinase inhibitors has resulted in an increase in the performance lifespan to the point where the life expectancy of patients today is approximately equivalent to that of the general population. Fast-forward a few years, this understanding has provided us the most extraordinary example of a targeted medicine targeting the dysregulated kinase activity indicated by this molecular alteration.

Balducci, Lodovico & Dolan, Dawn (2014)[13] It is not apparent whether the effect of this age has been attributable to worse biology or other age associated conditions, including competitive deaths, serious chance of treatment-related complications or contraindications to bone marrow curative transplantation. The problem is relevant since it is expected that with the ageing of the population, median age of CML and disease prevalence in older people will steadily increase. This editorial looks at the biodiversity, efficacy and tolerance of TKI in the elderly and the age-adjusted lifespan of CML patients since the advent of TKI after physiological evaluation of the age definition.About 50% of patients with CML are 66 years old or older. Age was poorly predicted for survival in early research and has been included in all staged models as such.The introduction of better-tolerated tyrosine kinase (TKI) inhibitors than prior treatments and which extend the survival of most patients who react indefinitely allow for a fresh look at advanced age and LMF interactions

Patricia Weinschenker Bollmann (2011) [14] The usage of imatinib has improved the prognostics and outcomes of chronic myeloid leukaemia patients significantly. However, the presence of sensitivity or intolerant mechanisms prevents the disease from being eradicated in some patients We are presenting a complete evaluation of the advancements in the treatment of chronic myeloidalleucemia.The finding of the Philadelphia chromosome in 1960 and of the BCR-ABL oncogene in 1984 led to the development of targeted treatment for chronic leukaemia in future years, thereby modifying the natural history of chronic myeloid leukaemia.In most patients with T315I mutation, tyrosinekinase second-generation inhibitors are efficacious.

Buyukasik, et al. (2010)[15]This is the first illness in which fusion protein focused treatment was applied. The disease of "first' is frequently referred to as Chronic Myeloid Leukemia (CML). It would be the first condition to use the name leukaemia. It has been shown to be the first neoplastic illness linked to recurring chromosomalabnormalities. Important points about CML with special attention to therapy and monitoring were summarised in this review. In the last decade, both therapy and follow-up on CML has been fundamentally modified by the introduction of BCR-ABltyro-sine kinase inhibitors, first imatinib and then dasatinib and nilotinib, into clinical practise.

Alfonso Quintás-Cardama(2006)[16]Since it was discovered that chromosomatic anomaly was distinct and continuous a little over 40 years ago, significant progress in the biology of this disease has been accomplished.The model for study and control among malignant illnesses has now become Chronic Myeloid leukaemia (CML). This progress has led to a considerable improvement in people with this disease's long-term prognosis. This alteration was first introduced by the utilization of stem cell transplantation and interferon alfa, but the era of molecular therapy has opened recently. The prospects today are significantly higher for CML patients than they were only a few years earlier. We are looking forward to repeat this exciting CML journey into other illnesses. We examine our present understanding of this condition in this article.The clearest illustration of our efforts to detect molecular anomalies and build specialised pharmacologies can be imatinib, a potent selective tyrosine-kinase inhibitor. Furthermore, it has led to the rapid creation of novel agents that can overcome the resistance by understanding at least some mechanisms of resistance to imatinib.

4. CATEGORIZATION ALGORITHMS AVAILABLE FOR CURING LEUKEMIA

A. Support Vector Machines: SVM is a categorical classifier, for classification in lymphoid and myeloid stem cell sample blood pictures as acute lymphoblasty and acutes myeloid leukaemia. SVM is a binary classification algorithm. The precision achieved using SVM, Laosai et al. (2014) was 92 per cent. In SVM a separating surface separates the input space of the data sets, optimising the class margin. A classifier training algorithm identifies the support vectors. Sample form and texture-based characteristics are retrieved and recorded for each nucleus. From all functions the most important features are chosen and used for training SVM. The selected features include number of nuclei lobes, the ratio between nuclei and cells, the perimeter ratio to nuclei and entropy. The quantity and arrangement of the nuclei lobe are the main criteria used to determine the WBC class[17]

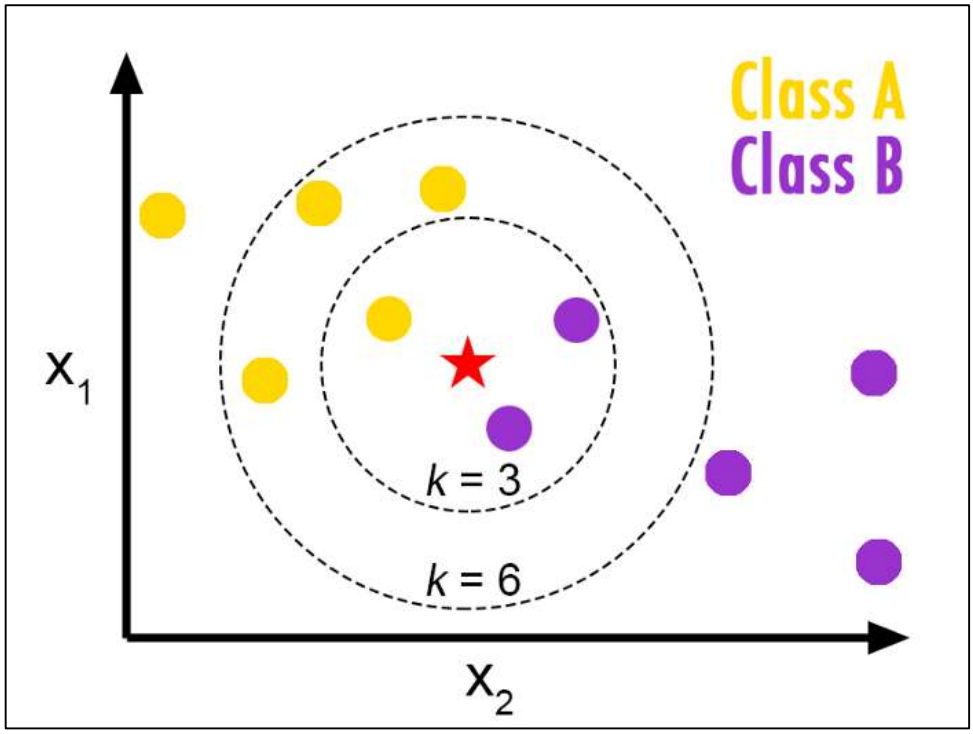

Figure 2: A simple KNN analyzer

B. K-Nearest Neighbour In order to classify the leukaemia in ordinary blood cells, Subhan et al. (2015) utilized KNN algorithms to show KNN as a good-scalability omnipresence classification method. The k-NN method categorises new items on the basis of similarity measurements assuming that comparable objects exist close by [18]. The KNN algorithm is a lazy learner with no phase of training and only storing his training data set, without taking up any responsibility of creating a model that has multiple parameters. SUPARDI et al. (2012) have also classified cells as Acute MyelogenicLeukemia (AML) and Acute Lymphocytic Leukemia (ALL) with a precision of 80% for blasts in leukaemic cells and have employed k-NN. Twelve basic characteristics of leucemic blood pictures have been studied on several occasions to depict size, colour and form and K values and distance metrics. The k-NN classification yielded good results with a frequency of k=4 and a cosine distance measure.[19].

C. Neural Networks Neural networks is used by Vincent et al. (2014) to classify images of blood smear into regular blood cells and leukaemia blood cells, therefore assisting in the diagnosis of leukaemia as a medical decision system. Preparation stages include preprocessing, grouping and segmentation of images. Following this Principal Component Analysis (PCA), the main input components that constitute part of the Neural Network Classifier are extracted [20]. This classification is intended to discriminate between normal (non-cancer) and abnormal cells (cancerous). The first two input layer nodes are supplied by PCA with the very first two main component outputs. The input load and a weight and a cached node bias are received and its input is obtained in accordance to Equation 1.

$$\boldsymbol{ni_j = \sum_{k=1}^{2}\sum_{j=1}^{3} xi_k . wi_{kj} + Bi_j} \qquad \textbf{(1)}$$

Equation 2 will be used to compute the observed result node value.

$$\boldsymbol{output = \sum_{j=1}^{3}\left(\sum_{k=1}^{2} xi_k * wi_{kj} + Bi_j\right) * Wo_j + B_j} \qquad \textbf{(2)}$$

This neural network's rudimentary structure is presented in figure 2

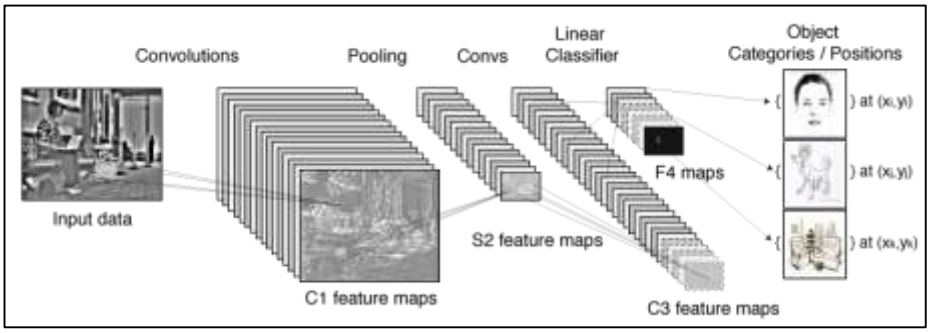

Figure 3: A simple diagram of CNN

D. Naïve Bayes : Gautam et al. (2016) used Naïve Bayes for the classification of leukocytes [21]. Bayes Theorem is an independent Naïve classifier that focuses on the analysis of the characteristics does not depend on whether other characteristics exist or do not and each of these characteristics contributes to the probability independently. Since relatively modest amounts of data are

required for training, many parameters required for categorization can be estimated. The estimation of its parameter is based on the maximum probability method as a supervised classifier. First, the average and variation of each of the characteristics of each cj class. It is expected that the chance of pre-p(c) is considered outcome variable and saved afterwards as

$$\boldsymbol{P(C_j) = \frac{Tc_j}{T_i}} \quad \textbf{(3)}$$

In the T_{CJ} class c, the number of trained images is T_i, with j=1,2,3,4,5 in the class c. Utilizing the Naïve Bayes classifier using Bayes rule as the back probability may be found:

$$\boldsymbol{P(C_j|x_1.x_{2,\ldots.}x_n) = P(x_1.x_{2,\ldots.}x_n|C_j) * P(C_j)} \quad \textbf{(4)}$$

Where $P(C_j|x_1.x_{2,\ldots.}x_n)$is the later chance of x belonging to c_j. Naïve Bayes assumes that the probability of likelyness can be determined through conditional probability of attributes as

$$\boldsymbol{P(x|C_j) = P(x_1C_j) * P(x_2C_j) * \ldots * P(x_nC_n)} \quad \textbf{(5)}$$

And the subsequent chance can now be determined as:

$$\boldsymbol{P(C_j|x_1.x_{2,\ldots.}x_n) = P(C_j) * P(x_1|c_j) * P(x_2|c_j) * \ldots * P(x_n|c_j)} \quad \textbf{(6)}$$

With an exactness of 80.88 percent each white blood cell is classified with highest reverse probabilities [21]

E. Deep Learning: Deep learning was employed for classifying blood imaging into ALL subtypes, i.e. L1, L2 and L3, or A. Rahman et al. Deep learning (2019). CNNs is one of the deep learning algorithms widely utilised in data processing of biological images. The manual extraction is excluded and characteristics are learned immediately from the image and then complemented by input data to provide a precise result for the classification. The classification process depends to some part on the discriminating characteristics because too many characteristics confuse the classifier and too few characteristics fail to provide efficient categorization. The data set, which was separated as training and test data, included blood cell pictures of both ALL-historic and normal individuals. For classification of ALL in its subtypes or as standard it employed the Alexnet model with CNN. The last three models' layers, softmax and classification layer are completely related according to the data [22]. The initial stage is to design the CNN architecture, and then to define the input layer and the convolutionary layer. The input layer sets the picture dimension to CNN that matches the width, the height and the number of image channels (single channel for grayscale image and 3 for RGB image). The convolution layer or second layer is made up of neurons connecting the sub-area of the image or the output layers preceding it. After scanning the picture, a convolutionary layer learns the characteristics of these locations. A normalising layer is available between CL and CLU, so that the training process can be speeded up and sensitivity reduced. Max pooling layer follows the convolutionary layer and is used to sample the surplus sample. The output layer consisting of the softmax function [22] is followed by all characteristics merged in the last fully linked layer. A. Rahman et al. (2019) employed a set of 100 L1, 100 L2, 30 L3 and 100 regular photographs. The CNN showed 97,78 per cent accuracy by taking 1,00e-04 as a learning rate and at 20 epochs with 80 per cent as training data and 20 per cent like test data. Figure 4 illustrates the CNN used.

5. PROPOSED METHODOLOGY

For instance, RF, SVM, k-NN, LDA, GBC and DT are six of the AI algorithms in this Article. We use this data on the subject from the data file on the (diagnostic) threat of *CML*. We have chosen these six algorithms to identify the best method for preventing the pace of disease with or without SFS. In the midst of them there was a correlation. A cross-approval system is used to prepare the classifier to learn. The output from basic students and output from the decreased SFS processes were used for the inspection of the data set in the last analysis. Figure 4 shows the model of construction for the job suggested.

- The information index is divided into training and testing.
- All data will be pre-processed and missing values will be replaced by training data set patterns and procedures.
- The train consists of ten sections.
- The main model fits the full information system of the train at this stage.
- Adapt to the 9-part basic model and create 10th component forecasts. It is applicable to all training sets.
- Predictions on the test set can be made using this model. Repeat steps iii to v for another fundamental model to get another prediction set for the group of trains and the test group.
- This model is used to forecast the test expectations, which extends the accuracy rate.
- The train figures are utilised for the construction of new models.

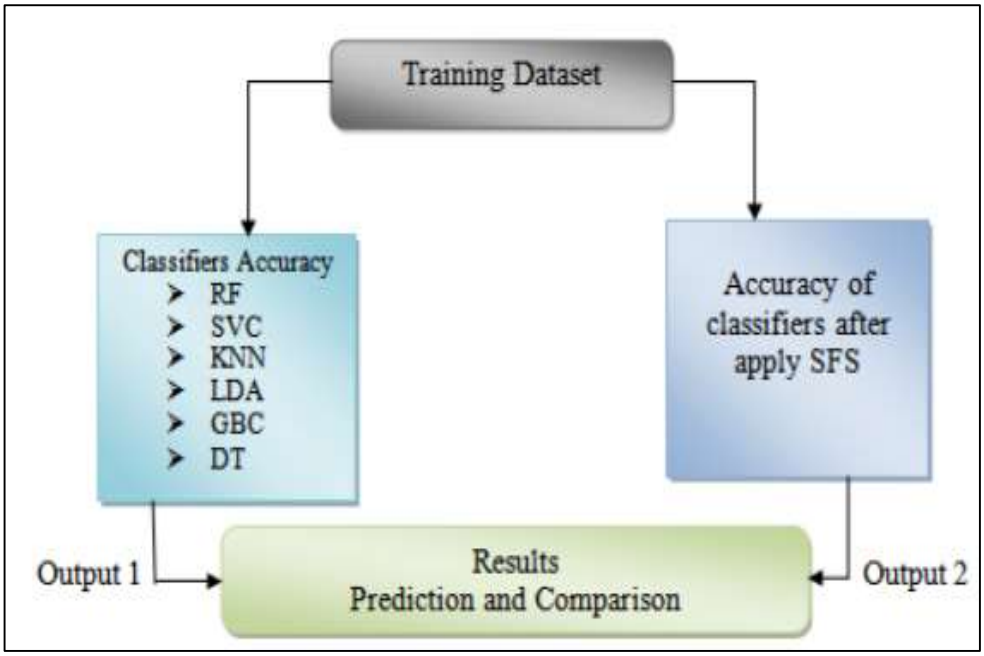

Figure 4: Proposed Model

6. EXPERIMENTAL SETUP

All the tests on a PC were carried out using Python's 3.6 stages, with a 2.4 GHz processor and 4 GB RAM. In order to evaluate the presentation, we use 10 times cross-validation to calculate five ML on each information indice. For all ML calculations we use the default limits of Python.

1. **Dataset**The standard UCI machine learning repository [23] Wisconsin (diagnostic) dataset was used for training and testing purposes. The database includes 569 instances and 32 functionalities and the table is not missing. Each characteristic is illustrated in Table 1 below.

Table 1: Information Attribute

Attribute	Attribute Information
ID number	Dropped (not relevant)
Diagnosis	2 Classes (B - 425, M- 128)
Radius	average of distances from center to points on the perimeter
texture	std of gray-scale values
area	Mean area of the nucleus
smoothness	local variation in radius lengths
compactness	given by formula "perimeter^2 / area - 1.0"
concavity	how severe are the concave portions of the contour
Concave points	number of concave portions in the contour
Symmetry	How symmetrical are the contours in the region
Fractal Dimension	Given by formula “coastline approximation - 1”

7. RESULTS

We conducted experiments using*CML*datasets in order to evaluate the performance of our approach. The correlation matrix is used as an input as an input for further analysis and for advanced analysis in order to summarise the data. The correlation matrix of the dataset for *CML* is shown in Figure 5 In order to detect less significant components; the principal component analysis is carried out on the information collection [24]. We have used PCA (as presented in figure 6) to validate that the variables are independent of each other.

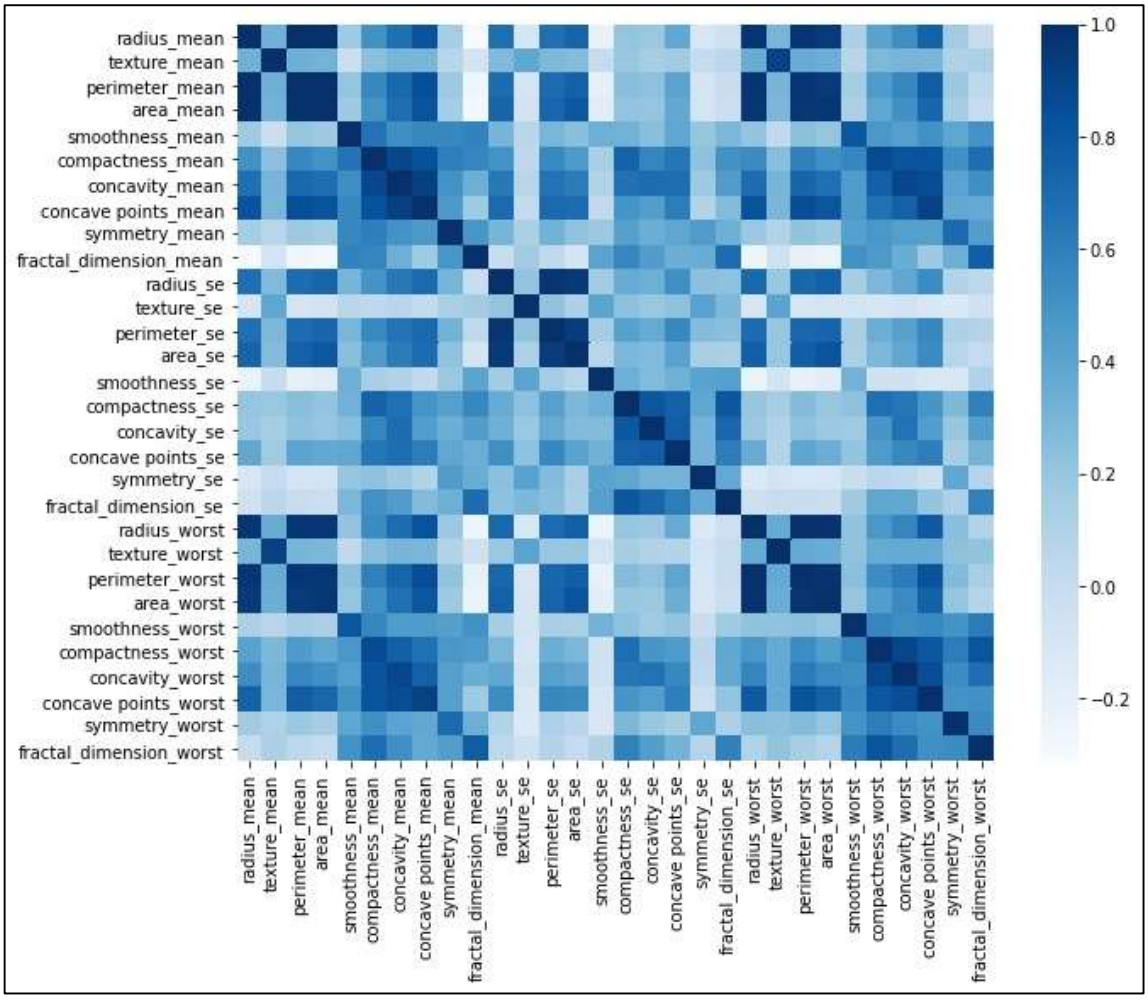

Figure 5: Dataset Correlation Matrix

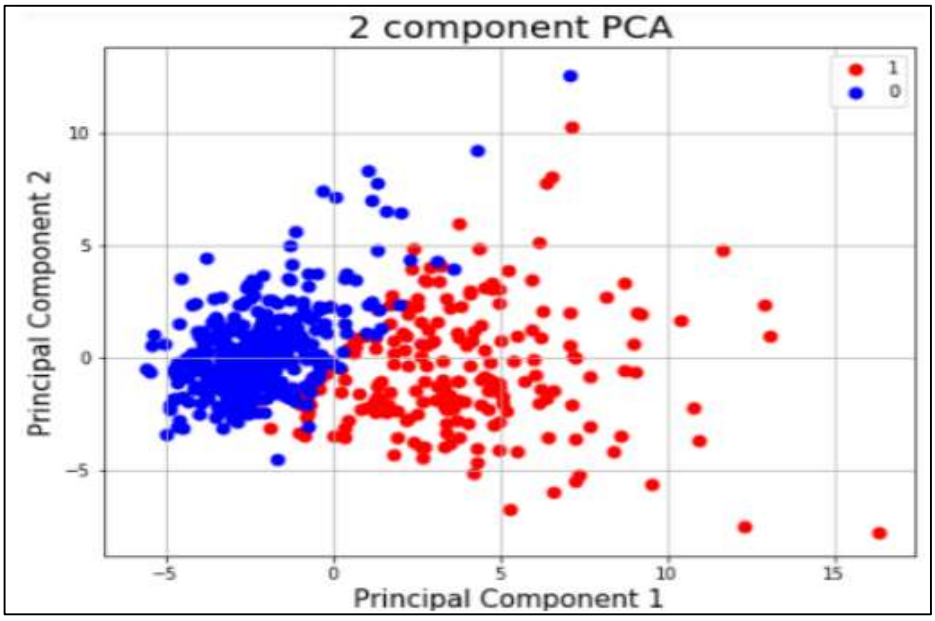

Figure 6: Main component analysis

$$\boldsymbol{Accuracy} = \frac{\boldsymbol{True\ Positive + True\ Negative}}{\boldsymbol{Total}} \quad (7)$$

Table 2 below lists the classifier's precise testing and training and its SFS accuracy.

Accuracy estimates the importance of each classifier [25]. The precision is determined on the basis of the amount of the rights categories that are profitable and risky to the number of instances, and is dictated by the use of the conditions: Figure 7 shows the statistical depiction of Table 2

Table 2: Classification accuracy with and without SF

Classifiers	Test_(Accuracy)	Train_(Accuracy)
RandomForestClassifier_sfs	96.67	**100.00**
KNeighborsClassifier_sfs	95.68	**95.26**
SVC_rbf	96.76	**97.25**
SVC_sfs	96.66	**97.50**
LinearDiscriminantAnalysis_	95.08	**96.24**
LinearDiscriminantAnalysis_sfs	96.91	**96.24**
RandomForestClassifier_	97.91	**98.76**
KNeighborsClassifier_	96.92	**98.48**
SVC_linear	94.92	**99.75**
GradientBoostingClassifier_	95.75	**100.00**
DecisionTreeClassifier_sfs	95.16	**100.00**
RandomForestClassifier_FS	94.56	**100.00**
DecisionTreeClassifier_	93.97	**100.00**
GradientBoostingClassifier_sfs	93.97	**98.76**
SVC_poly	**89.88**	**87.71**

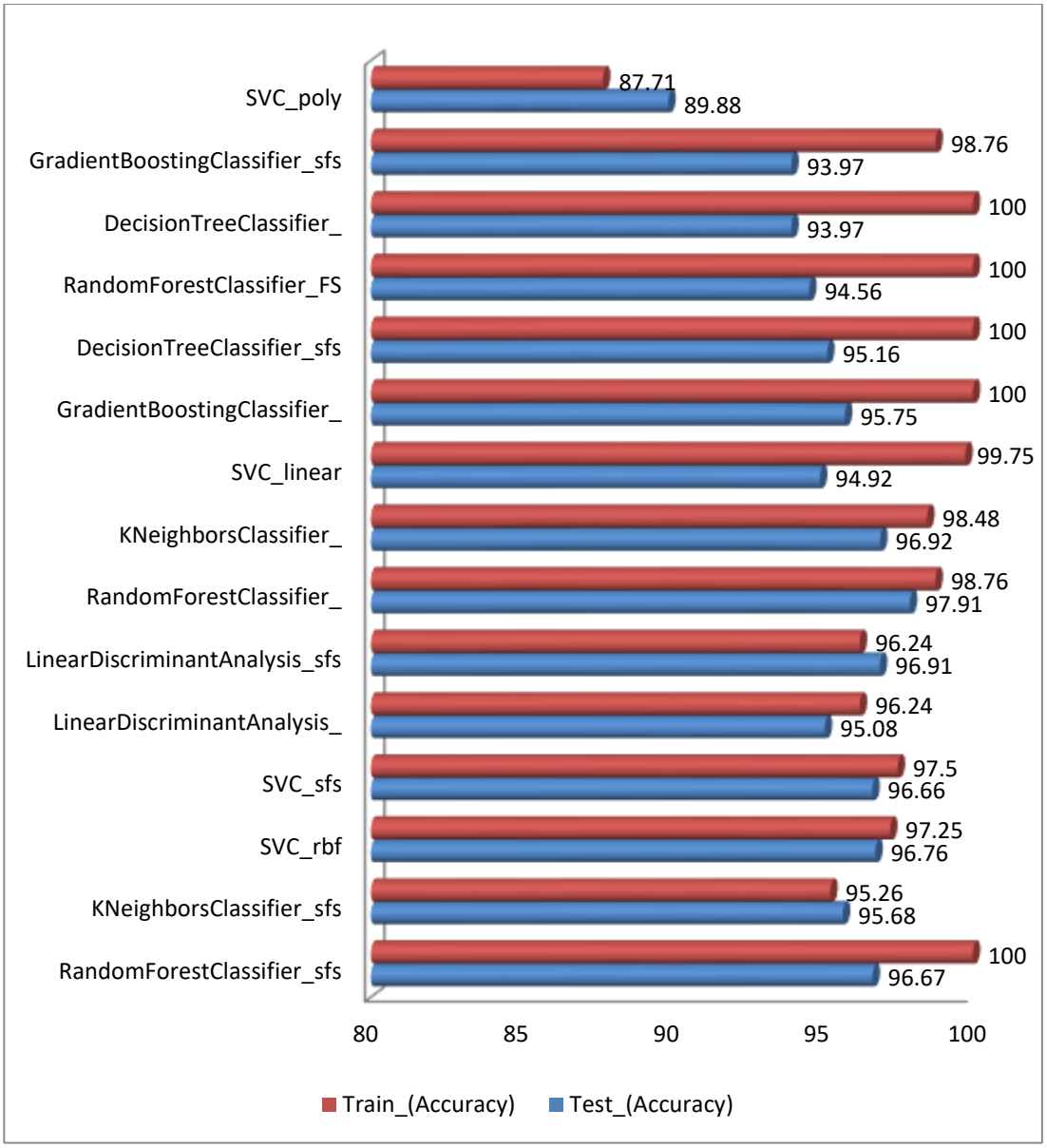

Figure 7: Precision of categorisation often repeated with SFS

8. DISCUSSION

The merits and de-merits of all algorithms are unique. Tables 2 illustrate that the algorithms are controlled by Naïve Bayes and SVM, and unattended ANN and CNN and k-NN are unattended algorithms. SVM is the only binary classification system, although all other methods can be classified in more than 2 classes. SVM and k-NN are good for smaller datasets, while Neural and Naïve Bayes function with both large and small datasets, and with a large dataset, deep education works fine [26, 27]. Naïve Bayes classifies linear data, while k-NN, neural networks and profound education work along with non-linear data, and SVM works well with non-linear as well as linear data. With respect to the precision settings set out in Table 2 it cannot be determined that CNNs are best because the evaluation is not based on same dataset. However, comparisons are done between different algorithms on specific criteria and always keep in mind the principle of Occam's Razor: use the least complicated and only complicated algorithm that meets the requirements if strictly necessary.[28]' The various algorithms presented in Table 3 are based on merits and demerits.

Table 3: Comparison of Classification Algorithms

Algorithm	Supervised/Unsupervised	Suitable classes count	Dataset type	Used data(Linear/Non-linear)	Accuracy
Support Vector Machine	Supervised	Binary	Small	Linear and Non Linear	93.0%
K- Nearest Neighbour	Unsupervised	Multiple	Small	Non Linear	81.0%
NN	Both	Multiple	Small and Large	Non Linear	93.8%
NB	Unsupervised	Multiple	Small and Large	Linear	80.09%
DL	Both	Multiple	Large	Non Linear	97.89%

Table 4: Advantages and Disadvantages

Algorithm	Advantages	Disadvantages
SVM	High precision Not essential linearly separable space Works well with unstructured and semi-structured data Scales well to high-dimensional information	Does not scale to big datasets Long duration of training Binary classification
K-NN	No required training periods Easy to apply No required training period	Cannot handle large datasets Cannot handle large datasets High cost of computing High cost of computing
Neural Networks	Works on large and small data sets efficiently Requires only little statistical training Detection of complex non-linear relationships between dependent and independent variables Faults tolerant	High computational overhead Substantial overfitting Network explanatory behaviour leads to issues unclear network duration
NaïveBayes	Simple Classifier Fast convergence Only less training information required Works efficiently both for big and small datasets.	Assumes all attributes are independent linearly; but that is not the case in real world. Chance of accuracy loss Unable to adjust addictions Assumes that there exist numerical qualities

For the classification of benign and malignant tumours, some information mining approaches are used. In this study, six calculations of supervised learning classification and the sequential segment research are used to forecast and dissect *CML*to various limitations. Here the main focus is which classifier is more precise. As a result, 97.66 per cent is the highest and most notable precision of RF sfs, KNN sfs, SVC rbf, and SVC sfs, and their residual accuracy is lower. Benevolent and 212 *CML*cases were each out of 257 cases from the information index above Table 3. We can assume that the identification of successive components is the optimal method to anticipate *CML*growth for all exploratory results based on accuracy and different limits.

9. CONCLUSION

Techniques for machine learning are becoming more popular, and this research attempts to compare five classification algorithms for machine learning: Vector support systems, KNearest, Neural Networks, Naïve Bavaria and Deep learning. Literature has been published on these *CML*and prediction systems. Further research will result in the effective implementation of appropriate Machine Learning algorithms for the identification of the effect of leukemic patient therapy.

REFERENCES

1. Cortes J.(2004) Natural history and staging of chronic myelogenous leukemia. Hematol Oncol Clin North Am.18(3):569-84
2. Melo JV, Deininger MW.(2004;) Biology of chronic myelogenous leukemia--signaling pathways of initiation and transformation. Hematol Oncol Clin North Am. 18(3):545-68, vii-viii.
3. Tefferi A, Dewald GW, Litzow ML, Cortes J, Mauro MJ, Talpaz M, et al. (2005) Chronic myeloid leukemia: current application of cytogenetics and molecular testing for diagnosis and treatment. Mayo Clin Proc.;80(3):390-402
4. Cortes JE, Talpaz M, OBrien S, Faderl S, Garcia-Manero G, Ferrajoli A, et al.(2006)Staging of chronic myeloid leukemia in the imatinib era: an evaluation of the World Health Organization proposal. Cancer. 106(6):1306-15.
5. Kantarjian H, Schiffer C, Jones D, Cortes J. (2008) Monitoring the response and course of chronic myeloid leukemia in the modern era of BCR-ABL tyrosine kinase inhibitors: practical advice on the use and interpretation of monitoring methods. Blood. 111(4):1774-80
6. O'Hare T, Eide CA, Deininger MW(2007) Bcr-Abl kinase domain mutations, drug resistance, and the road to a cure for chronic myeloid leukemia. Blood.;110(7):2242-9
7. Meyn MA 3rd, Wilson MB, Abdi FA, Fahey N, Schiavone AP, Wu J, et al. (2006)Src family kinases phosphorylate the Bcr-Abl SH3-SH2 region and modulate BcrAbl transforming activity. J Biol Chem.;281(41):30907-16.
8. Shah NP, Tran C, Lee FY, Chen P, Norris D, Sawyers CL.(2004) Overriding imatinib resistance with a novel ABL kinase inhibitor. Science.305 (5682):399-401.
9. Cortes, J., Lang, F. (2021) Third-line therapy for chronic myeloid leukemia: current status and future directions. J Hematol Oncol 14, 44 https://doi.org/10.1186/s13045-021-01055-9
10. Hehlmann, Rüdiger. (2020). Chronic Myeloid Leukemia in 2020 HemaSphere. 4. e468. 10.1097/HS9.0000000000000468.
11. Bonifacio, Massimiliano & Stagno, Fabio & Scaffidi, Luigi & Krampera, Mauro & Di Raimondo, Francesco. (2019). Management of Chronic Myeloid Leukemia in Advanced Phase. Frontiers in Oncology. 9. 1132. 10.3389/fonc.2019.01132.
12. Thompson et al (2015)" Diagnosis and Treatment of Chronic Myeloid Leukemia (CML) in 2015" Mayo Clin Proc. 90(10): 1440–1454
13. Balducci, Lodovico & Dolan, Dawn. (2014). Chronic Myelogenous Leukemia (CML) in the elderly. Mediterranean journal of hematology and infectious diseases. 6. e2014037. 10.4084/MJHID.2014.037.
14. Patricia Weinschenker Bollmann (2011)" Chronic myeloid leukemia: past, present, future" einstein. 2011; 9(2 Pt 1):236-43
15. Buyukasik, Yahya &Haznedaroglu, Ibrahim & Ilhan, Osman & Grp, Turkish. (2010). Chronic Myeloid Leukemia: Practical Issues in Diagnosis, Treatment and Follow-Up. Uhod-UluslararasiHematoloji-OnkolojiDergisi. 20. 1-12.
16. Alfonso Quintás-Cardama (2006) "Chronic Myeloid Leukemia: Diagnosis and Treatment" Mayo Clin Proc. 81(7):973-988
17. J. Laosai and K. Chamnongthai. "(2014) Acute leukemia classification by using SVM and K-Means clustering" Proceedings of the International Electrical Engineering Congress pp. 1-4.
18. Subhan, Ms. Parminder Kaur. (2015) Significant Analysis of Leukemic Cells Extraction and Detection Using KNN and Hough Transform Algorithm‖ International Journal of Computer Science Trends and Technology, vol. 3, no.1, pp. 27-33,
19. Supardi, N. Z., Mashor, M. Y., Harun, N. H., Bakri, F. A., & Hassan, R (2012). ―Classification of blasts in acute leukemia blood samples using k-nearest neighbour‖. IEEE 8th International Colloquium on Signal Processing and Its Applications, pp. 461-65,
20. Vincent, I., Kwon, K.-R., Lee, S.-H., & Moon, K.-S. (2015). ―Acute lymphoid leukemia classification using two-step neural network classifier‖ 21st Korea-Japan Joint Workshop on Frontiers of Computer Vision (FCV) Jan. 2015

21. Gautam, A., Singh, P., Raman, B., &Bhadauria, H.(2016) ―Automatic classification of leukocytes using morphological features and Naïve Bayes classifier‖ IEEE Region 10 Conference (TENCON), pp.1023-27,

22. Rehman, A., Abbas, N., Saba, T., Rahman, S. I. ur, Mehmood, Z., &Kolivand, H. (2018)―Classification of acute lymphoblastic leukemia using deep learning‖ Microscopy Research and Technique, vol. 81, no. 11,

23. UCI Machine Learning Repository. [http://archive.ics.uci.edu/ml/]. Irvine, CA: University of California, Center for Machine Learning and Intelligent Systems.

24. Nakazato, Y., Sugiyama, T., Ohno, R., Shimoyama, H., Leung, D.L., Cohen, A.A., Kurane, R., Hirose, S., Watanabe, A. and Shimoyama, H., (2020). Estimation of homeostatic dysregulation and frailty using biomarker variability: a principal component analysis of hemodialysis patients. Scientific Reports, 10(1), pp.1-12.

25. Chicco, D. and Jurman, G., (2020). The advantages of the Matthews correlation coefficient (MCC) over F1 score and accuracy in binary classification evaluation. BMC genomics, 21(1), p.6.

26. Chaurasia, V. and Pal, S., (2020). Applications of machine learning techniques to predict diagnostic breast cancer. SN Computer Science, 1(5), pp.1-11.

27. Chaurasia, V., Pandey, M.K. and Pal, S., 2021, March. Prediction of Presence of Breast Cancer Disease in the Patient using Machine Learning Algorithms and SFS. In IOP Conference Series: Materials Science and Engineering (Vol. 1099, No. 1, p. 012003). IOP Publishing.

28. Aggrawal, R. and Pal, S., (2020). Sequential Feature Selection and Machine Learning AlgorithmBased Patient's Death Events Prediction and Diagnosis in Heart Disease. SN Computer Science, 1(6), pp.1-16