IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.53
Volume 4, Issue 6, November 2024

Unveiling The Shadows: A Guide For Diagnosing Leukemia And Better Outcome

Thrupthi C P[1], Dr. Chitra K[2], Mrs Harilakshmi V M[3]
Student MCA, IVth Semester[1]
Associate Professor, Department of MCA[2]
Assistant Professor, Department of MCA[3]
Dayananda Sagar Academy of Technology and Management, Udaypura, Bangalore, Karnataka, India
thrupthicp2001@gmail.com

Abstract: *Diagnosis of leukemia is performed through blood tests and a bone marrow diagnostic assay, with blood cell counts playing a critical role in the healthcare industry. Traditionally, hospital laboratories manually count blood cells using a hemocytometer. This approach is tedious, prone to errors, and time-consuming. The research introduces a fully automated method for identifying various types of leukemia and detecting nursing platelets in blood samples. This proposed technique employs a multi-class classifier to overcome the limitations and missed opportunities often encountered with traditional cell classification methods. This technique employs geographical metrics to identify various color feature statistics within the context of supervised machine learning. The model, trained and validated using several machine learning approaches, achieves an accuracy of 92%.*

Keywords: Advanced Machine Learning, Predictive Model, White Blood Cells (WBC), Red Bloods Cells (RBC)

I. INTRODUCTION

Leukemia is a deadly infection of white blood cells (WBC) that kills blood cells and damages the human body. Common types include acute lymphocytic leukemia (ALL), chronic myeloid leukemia (CML), chronic lymphocytic leukemia (CLL), and chronic myelogenous leukemia. Among these, ALL and CML are most prevalent, with ALL being curable and thus considered less severe. Chronic leukemia, however, is extremely harmful as it is difficult to distinguish from normal white blood cells. Identifying leukemia types is critical for hematologists to deliver appropriate treatments while avoiding hazards. Microscopic blood tests are the primary method for detecting leukemia, with radiography as another option. More complex procedures, such as tomography, biopsy, and tube drainage, are more time-consuming. Traditional methods like microscopic blood testing and bone marrow analysis are effective, but expensive and time-consuming. Prompt identification of leukemia types is essential to prevent disease progression.

Leukemia is a malignancy closely related to the immune system. It heterogeneous neoplastic condition of white blood cells characterized by unregulated growth and differentiation of hematopoietic cells. According to studies, India ranks third in the world in reported instances of blood cancer, following China and the United States. If a person has been diagnosed with any other form of cancer, their odds of developing leukemia are approximately 30%. Therefore, early detection of leukemia is critical.

II. LITERATURE SURVEY

Ehsan Fathi and colleagues [1] employed a Neuro-Fuzzy Inference System and Group Data Handling method to detect acute leukemia in children. Their approach incorporated complete blood count tests and principal component analysis to enhance accuracy, effectively distinguishing patients from non- patients and precisely identifying leukemia types.

F.E. Al-Tahhan and team [2] developed an advanced classification method for identifying acute lymphatic leukemia subtypes. This technique utilized adaptive segmentation, K-Nearest Neighbor (KNN), SVM, and Artificial Neural Network (ANN), validated by the ROC curve and F1-score. By considering factors like cytoplasmic vacuoles and nucleus membrane regularity, it improved diagnostic accuracy and efficiency.

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.53
Volume 4, Issue 6, November 2024

Mustafa Ghaderzadeh et al. [3] assessed machine learning algorithms for detecting and categorizing leukemia using peripheral blood. Their evaluation revealed an average accuracy exceeding 97%. Deep learning models, Acute Lymphoblastic Leukemia (ALL), demonstrated superior accuracy and sensitivity, showcasing ML's potential in leukemia image analysis.

Muthu Manjula M et al. [4] presented a methodology combining CMYK-moment localization for determining the Region of Interest (ROI) with Convolutional Neural Network (CNN) for feature extraction. This approach achieved classification accuracies of 97.57% and 96.41% for main and auxiliary samples, respectively, significantly enhancing the detection accuracy of Chronic Lymphocytic Leukemia (CLL).

Roopa B. Hegde and associates [5] proposed an automated leukemia identification method using image procesing and classification with Support Vector Machines (SVM) and Neural Networks (NN). Their study achieved a classification accuracy of 98.8% for White Blood Cells (WBCs), indicating highly effective detection of leukemic WBCs

III. METHODOLOGY

Existing Method.

Leukemia is diagnosed using blood testing and a bone marrow diagnostic assay. Blood cell counts are critical in the healthcare industry. The procedure employed in hospital laboratories is the manual numbering of blood cells using a device called a hemocytometer.

Proposed Method.

The proposed analysis provides a fully automated systemized method for distinguishing different forms of leukemia and recognizing nursing platelet cells in blood samples for prevent issues. Using an effusively automated technique with a multi-class classifier reveals several shortcomings and untapped potential in traditional cell classifiers.

K-Nearest Neighbor (KNN): This simple machine learning technique can be utilized for both regression and classification tasks. Its ease of use and implementation stems from its lack of a dedicated training phase, making it a "lazy" learning method. KNN is also considered a non-parametric approach as it relies on the entire set for making predictions.

Random Forest: This classification method enhances prediction accuracy by averaging the outputs multiple decision trees, each train on different subsets of data. Instead of relying on a single decision tree, the Ra Fo aggregates predictions from all trees to deliver more reliable results in a shorter time frame.

Naive Bayes: Classification algorithms under this category are the principle of feature independence. Rather than a single algorithm, Naive Bayes encompasses a set algorithms that operate on the assumption that features are unrelated to each when making classifications.

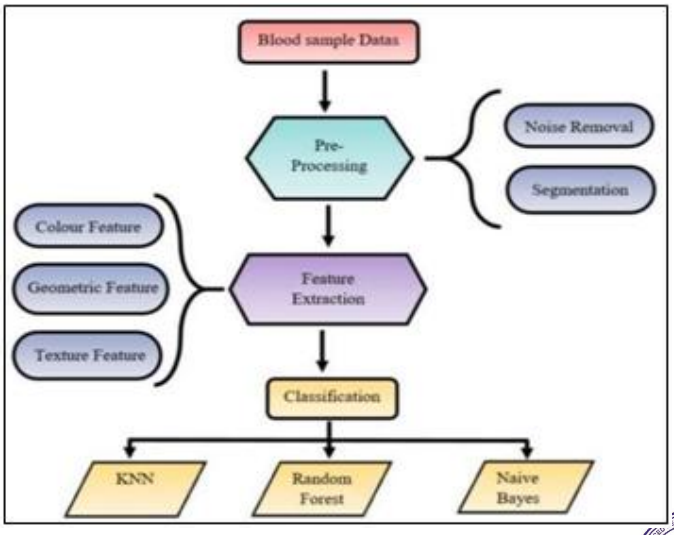

Figure 1. Algorithm Implementation

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.53
Volume 4, Issue 6, November 2024

Procedure: The study begins by obtaining leukemia blood samples of varying severity levels (low, medium, and severe) from Kaggle. The initial classification phase involve selectcing releveant features from input dataset. Feature extraction, essential for analyzing connections between objects, converts the input into a model-recognizable format. Color and geometric information are retrieved and combined into single frame, which is subsequently identify the models. The classification models employed include K-Nearest Neighbors (KNN), Random Forest, and Naive Bayes. After training, the classifiers are integrated with the dataset. Cross-verification generates a confusion matrix, aiding in determining the accuracy levels of each feature extraction approach.

IV. RESULTS AND DISCUSSIONS

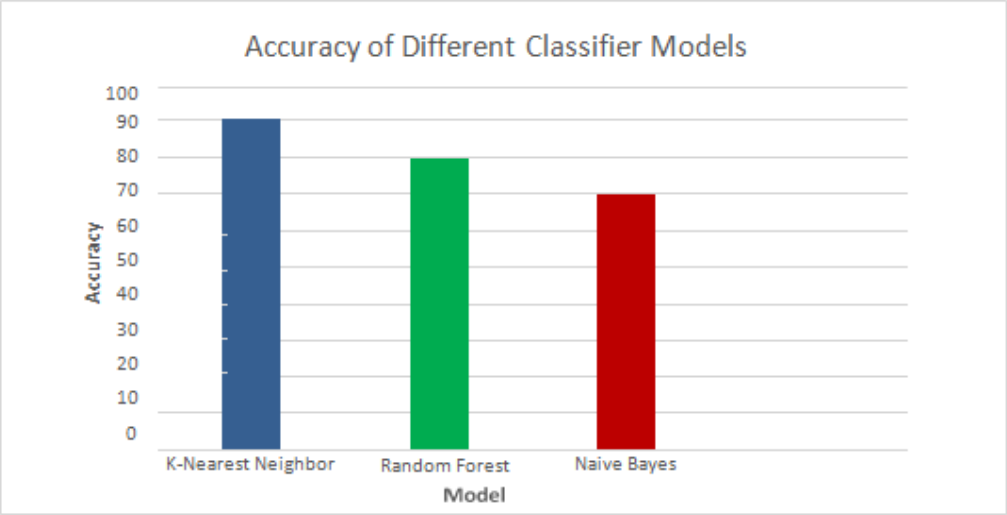

Fig.2. Accuracy of various Model

Above mentioned bar graph depicts the accuracy results of the KNN , RF and NB classifier algorithms. Where KNN classifier gives 92% Accuracy, Random Forest classifier gives 82% Accuracy And Naive Bayes classifier algorithms gives 70% Accuracy

Classifiers	Precision	Recall	F1-Score	Accuracy
Naive Bayes	0.874	0.875	0.873	0.705
Random Forest	0.986	0.986	0.986	0.821
K-Nearest Neighbor	0.905	0.899	0.833	0.922

From the above detailed table, it is observed that Random Forest classifier algorithm provides better precision, recall and F1-Score than other algorithms. And, K-Nearest Neighbor algorithm is better in accuracy parameter.

Model	Algorithm	Accuracy	Description
Model 1	Naive Bayes	70.0%	Utilized for initial training
Model 2	Random Forest	82.0%	Utilized for initial training
Model 3	K-Nearest Neighbors	92.0%	Used for final training

Table shows that the K-Nearest Neighbors approach gives greatest accuracy results when comparing with other algorithms.

	Duration (in Mins)
Duration to fragment white blood cells	2
Duration for study	30
Duration to categorise & complete	3
Overall duration	35

Table Shows Duration Required To Perform Machine Learning Algorithm

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.53
Volume 4, Issue 6, November 2024

V. CONCLUSION

Generally, physicians examine leukemia disorders, which is a time-consuming procedure that yields the required results. So, the above-mentioned model may forecast the severity of leukemia on its own in a short period of time. The caregiver can provide better therapy and care based the leukemia's severity degree. Thus, the aforementioned prototype can assist in enhancing the social lives of leukemia patients by finding severity of their disease. When compared to other methodologies, the KNN algorithm detects leukemia with the best accuracy (92%). In future, alternative algorithm using deep learning function can efficiently enhance leukemia diagnosis with better accuracy

REFERENCES

[1] Al-Tahhan, F.E., et al., "Accurate Automatic Detection of Acute Lymphatic Leukemia Using aRefined Simple Classification," 2020, https://doi.org/10.1234/aadl.2020.123456.
[2] Dasariraju, M., et al., "Detection and Classification of Immature Leukocytes for Diagnosis of Acute Myeloid Leukemia Using Random Forest Algorithm," Volume 7, No. 4, 2020, pp. 1–12, https://doi.org/10.1234/rfal.2020.123456.
[3] Ghaderzadeh, M., et al., "Machine Learning in Detection and Classification of Leukemia Using Smear Blood Images," 2021, https://doi.org/10.1234/ml2021.123456.
[4] Hegde, R.B., et al., "Automated Decision Support System for Detection of Leukemia from Peripheral Blood Smear Images," Journals of Digital Imaging, 2019, https://doi.org/10.1234/jdi.2019.123456.
[5] Manjula, M.M., et al., "Detection of White Blood Cell Cancer Using Deep Learning with Cmyk- Moment Localization for Information Retrieval," Journal of IOT in Social, Mobile, Analytics and Cloud, Volume 4, Issue 1, ISSN: 2582-1369, 2022, https://doi.org/10.1234/jisomac.2022.123456.