CHRIST
(DEEMED TO BE UNIVERSITY)
BANGALORE · INDIA

A Machine Learning Framework for Leukemia Detection

Student Details:
Ben Dias Akkara (2262209)
Sandeep S (2262253)

Guided By: Prof. Chaithra P C

Agenda

- Introduction
- Methodology/Workflow
- Implementation Details
- Results and Discussions
- Conclusion
- Future Outlook

Introduction

Leukemia is a cancer of the blood and bone marrow that leads to abnormal production of white blood cells.

- **Detection Methods**:
  - **Complete Blood Count (CBC):** Identifies abnormal WBC, RBC, hemoglobin, and platelet levels.
  - **Bone Marrow Biopsy:** Examines bone marrow cells microscopically for malignant changes.
  - **Flow Cytometry & Immunohistochemistry:** Detects specific cell markers to differentiate leukemia subtypes.
  - **Genetic Tests (Cytogenetics, PCR, FISH):** Identify chromosomal abnormalities and gene mutations.

Types of Leukemia & Their Detection:

- **Acute Lymphoblastic Leukemia (ALL):** Detected through CBC showing high lymphoblast count; confirmed by bone marrow biopsy and flow cytometry.
- **Acute Myeloid Leukemia (AML):** Identified by presence of myeloblasts in blood/marrow; cytogenetic testing helps in classification.
- **Chronic Myeloid Leukemia (CML):** Detected by CBC (very high WBC count), bone marrow exam, and genetic testing for the **Philadelphia chromosome (BCR-ABL fusion gene)**.
- **Chronic Lymphocytic Leukemia (CLL):** Diagnosed mainly via CBC showing abnormal lymphocytes, confirmed with flow cytometry for cell surface markers.

Methodology/Workflow

- **Step 1 – CBC Analysis**
  - Collect Complete Blood Count (CBC) data from the patient.
  - Check for abnormalities such as abnormal WBC, RBC, Platelets, or Hemoglobin levels.
- **Step 2 – Bone Marrow Biopsy Analysis**
  - If CBC is abnormal, proceed with bone marrow biopsy (images and text reports).
  - Extract features from biopsy images and reports for further analysis.

- **Step 3 – Cancer Detection**
  - Use AI models to determine whether blood cancer is present
- **Step 4 – Leukemia Classification**
  - If cancer detected, classify into leukemia subtypes:
    - Acute Lymphoblastic Leukemia (ALL)
    - Acute Myeloid Leukemia (AML)
    - Chronic Lymphocytic Leukemia (CLL)
    - Chronic Myeloid Leukemia (CML)
- **Step 5 – Additional Diagnostic Tests**
  - Use supporting lab techniques like **Flow Cytometry** and **Immunohistochemistry** to improve classification accuracy.

Implementation Details

- The project was implemented in Python using libraries such as Pandas and Scikit-learn.
- The dataset (cbc information.xlsx) was processed by selecting only relevant CBC parameters.
- Each record was labeled as Normal or Abnormal based on standard medical reference ranges.
- A Random Forest Classifier was trained on the processed data for classification.
- The model was evaluated using accuracy, precision, recall, and F1-score to measure performance.

Results and Discussions

- The Random Forest model achieved high accuracy in classifying CBC results as Normal or Abnormal.
- The classification report showed strong precision, recall, and F1-scores for both classes.
- Since the labels were based on clear medical reference ranges, the model performance was very reliable.
- The system can assist healthcare professionals by quickly flagging abnormal test results for further review.

Conclusion

- The Random Forest model achieved high accuracy in classifying CBC results as Normal or Abnormal.
- The classification report showed strong precision, recall, and F1-scores for both classes.
- Since the labels were based on clear medical reference ranges, the model performance was very reliable.
- The system can assist healthcare professionals by quickly flagging abnormal test results for further review.

Future Outlook

•Incorporate larger biopsy image datasets for training deep learning models.

•Explore advanced image segmentation for bone marrow slides.

•Develop clinician-friendly dashboards for combined CBC + biopsy + test analysis.