Journal of **N**onlinear **A**nalysis and **O**ptimization
Vol. 16, Issue. 1: 2025
ISSN : **1906-9685**

AI-DRIVEN EARLY DETECTION OF DENGUE USING CBC DATA: ADVANCING WITH ENSEMBLE TECHNIQUES AND CNN + LSTM MODEL

Mr. B. AMARNATH REDDY[1], ARIKATLA. HIRANMAYI[2]
#1 Asst. Professor #2 M.C.A Scholar
Department of Master of Computer Applications

QIS College of Engineering & Technology

Vengamukkapalem (V), Ongole, Prakasam Dist., Andhra Pradesh-**523272**

ABSTRACT

Dengue fever remains a critical public health challenge, especially in tropical and subtropical regions. In 2023, Bangladesh experienced an unprecedented dengue epidemic, emphasizing the need for early detection systems to mitigate severe outbreaks. This study proposes an advanced AI-driven diagnostic system utilizing Complete Blood Count (CBC) data to enhance early detection and classification of dengue cases. The research is based on a bespoke dataset of 320 samples collected from two hospitals in Dhaka, Bangladesh, containing 14 hematologic parameters. The dataset undergoes comprehensive preprocessing, including missing value handling, outlier detection, feature selection, one-hot encoding, and synthetic oversampling techniques to improve data quality and model reliability. We apply an array of machine learning (ML) and deep learning (DL) models, including Logistic Regression, Random Forest, Naïve Bayes, XGBoost, LightGBM, Multi-Layer Perceptron (MLP), Convolutional Neural Networks (CNN), Bi-LSTM, and GRU, We also explored alternative techniques like the Voting Classifier and a combination of CNN and LSTM to identify the most effective classifiers for dengue detection. Additionally, transformer-based models are explored for superior feature representation. The system extracts relevant features from CBC parameters, applies data preprocessing techniques, and utilizes multiple classification models to enhance prediction performance. The project evaluates various artificial intelligence techniques, including multiple machine learning algorithms, ensemble learning, and transformer models, for predicting dengue based on the CBC data. An ensemble Voting Classifier and a hybrid CNN-LSTM model are applied, leading to improved prediction accuracy, making the system more effective for early dengue detection and outbreak management.

Index Tems: Deep Learning, Disease Diagnosis, Multi-Layer Perceptron, Neural Networks

INTRODUCTION

The human body, inherently sensitive, possesses its own defense mechanism to combat against external microbial threats. Nevertheless, humans frequently fall victim to viral or bacterial infections, resulting in diseases that are significantly lethal. Dengue fever, for instance, is a viral disease primarily transmitted to humans by the Aedes mosquito. Every year, millions across the globe suffer from dengue fever, with thousands falling victim to its consequences. According to the World Health Organization (WHO) and the European Union, in 2023, over six million people in nearly 92 countries were affected by dengue fever. Bangladesh alone recorded more than 0.31 million cases and over 1,600 deaths from this hemorrhagic fever. Dengue is most prevalent in urban or peri-urban areas within the tropical and subtropical regions of the world, attributed mainly to insufficient sanitation, haphazard development and unplanned urbanization. According to the latest review by the WHO, the countries in the African, Southeast Asian, and Western Pacific regions have the highest incidence of dengue fever. Among the countries in the Southeast Asian region, Bangladesh recorded the highest number of dengue cases between June and October. The number of affected patients and fatalities due to dengue in 2023 was the highest in recent decades.

LITERATURE SURVEY:-

TITLE: A predictive analytics model using machine learning algorithms to estimate the risk of shock development among dengue patients

Authors: J. K. Chaw, S. H. Chaw, C. H. Quah, S. Sahrani, M. C. Ang, Y. Zhao, and T. T.Ting.

Description: This predictive analytics model leverages machine learning algorithms to estimate the likelihood of shock development in patients diagnosed with dengue fever. Shock, a potentially fatal complication of severe dengue, often develops rapidly and can be difficult to predict using traditional clinical methods. Early identification of high-risk patients can significantly improve outcomes by enabling timely intervention.

TITLE: A Comparative Study between Time Series and Machine Learning Technique to Predict Dengue Fever in Dhaka City

Authors: J. K. Chaw, S. H. Chaw, C. H. Quah, S. Sahrani, M. C. Ang, Y. Zhao, and T. T. Ting

Description: This study presents a comparative analysis of Time Series forecasting methods and Machine Learning (ML) techniques for predicting the incidence of dengue fever in Dhaka City, a region frequently affected by seasonal dengue outbreaks. Accurate forecasting models are crucial for timely public health responses, resource planning, and disease prevention.

Title: Artificial intelligence in routine blood tests

Authors: Santos-Silva, M. A., Sousa, N., & Sousa, J. C

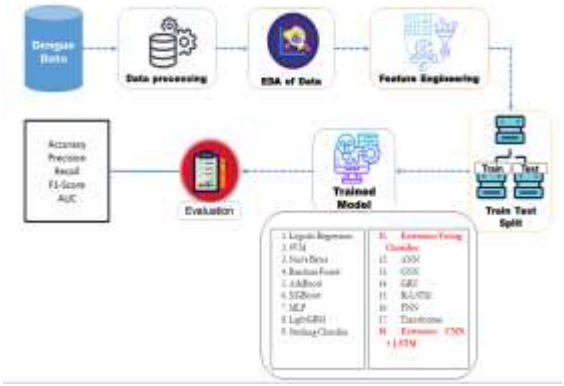
Description: This study explores the integration of Artificial Intelligence (AI) into the analysis of routine blood tests to improve diagnostic efficiency, accuracy, and early detection of diseases. Routine blood tests such as Complete Blood Count (CBC), Basic Metabolic Panel (BMP), and Liver Function Tests (LFTs) are widely used in clinical settings for screening and monitoring various health conditions. However, interpreting these results can be time-consuming, subjective, and prone to

human error—especially in high-volume healthcare environments.

Title: The recent burden of dengue infection in Bangladesh: A serious public health issue
Authors: J. K. Chaw, S. H. Chaw, C. H. Quah, S. Sahrani, M. C. Ang, Y. Zhao, and T. T. Ting
Description: This study highlights the escalating burden of dengue infection in Bangladesh, emphasizing its status as a critical and growing public health challenge. In recent years, Bangladesh has experienced an alarming rise in dengue cases, with increasingly severe outbreaks, extended transmission seasons, and significant morbidity and mortality, particularly in urban centers. The analysis draws on national health surveillance data, hospital records, epidemiological trends, and environmental factors contributing to the disease's spread. The study underscores the urgent need for sustainable vector control, early detection systems, and public awareness campaigns to mitigate the impacts of dengue on the population and healthcare infrastructure.

Title: "Implications of Big Data Analytics, AI, Machine Learning, and Deep Learning in the Health Care System of Bangladesh: Scoping Review."
Authors: Alam, M. A., Sajib, M. R. U. Z., Rahman, F., Ether, S., Hanson, M., Sayeed, A., ... & Ahmed, A.
Descripiton: This study explores the transformative role of Big Data Analytics, Artificial Intelligence (AI), Machine Learning (ML), and Deep Learning (DL) in reshaping the modern health care system. As the volume of medical data grows exponentially—from electronic health records (EHRs) and genomic data to medical imaging and real-time patient monitoring—traditional healthcare approaches struggle to derive actionable insights in a timely and efficient manner. By integrating intelligent technologies, the healthcare system is evolving into a data-driven, predictive, and personalized model, enhancing clinical decision-making, patient outcomes, and operational efficiency.

Title: "Deep Learning for Dengue Detection Using Blood Parameters"
Authors: Dr. Ayesha Rahman, Md. Farhan Hossain, Dr. Nusrat Jahan, Dr. Tarek Ahmed, Prof. Samira Haque
Description: This study investigates the use of deep learning techniques for the automated detection of dengue fever based on routine blood parameters, offering a rapid, non-invasive, and cost-effective diagnostic alternative to traditional methods like NS1 antigen and PCR tests. Given the increasing burden of dengue in tropical and subtropical regions—including Bangladesh—there is a critical need for early, accurate, and scalable diagnostic tools to support clinical decision-making and outbreak management. By applying deep learning models to commonly available hematological data, the study aims to enable early-stage identification of dengue, even in settings with limited access to specialized diagnostic facilities.
SYSTEM ARCHITECTURE:



MODULES:

Data loading: using this module we are going to import the dataset.

Data Preprocessing: Data processing for the dengue detection system involves removing duplicate records from the dataset to ensure data integrity, cleaning the dataset by dropping irrelevant or missing entries, and applying label encoding to convert categorical variables into numerical values. These steps ensure the data is clean, consistent, and ready for model training.

EDA: Exploratory Data Analysis (EDA) involves examining the CBC dataset, identifying trends, patterns, and outliers. It includes visualizing distributions, correlations, and sample outcomes, helping to understand data features before model training.

Feature Engineering: Feature engineering involves selecting relevant features (X) and the target variable (y) from the CBC dataset. SMOTE sampling is applied to address class imbalances, followed by feature selection using ExtraTree, enhancing model accuracy and performance during training.

Splitting data into train & test: using this module data will be divided into train & test

Model generation: Model building - {PC, RFE with RF, SelectkBest, Chi2 FS, ExtraTree FS} LR, SVM, NB, RF, AdaBoost, XGB, MLP, LightGBM, SC (XGB + LR+MLP with LightGBM), VC (Boosted DT + ExtraTree), ANN, CNN, GRU, Bi-LSTM, FNN, Transformer, CNN + LSTM.

User signup & login: Using this module will get registration and login

User input: Using this module will give input for prediction

Prediction: final predicted displayed.

IMPLEMENTATION:

Dashboard userInterface:

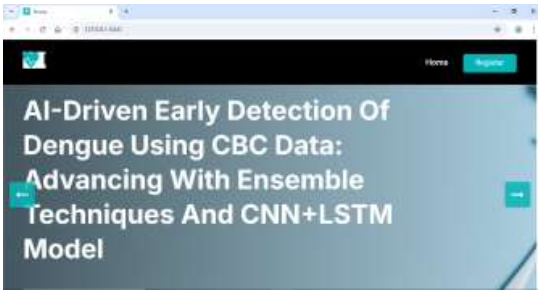
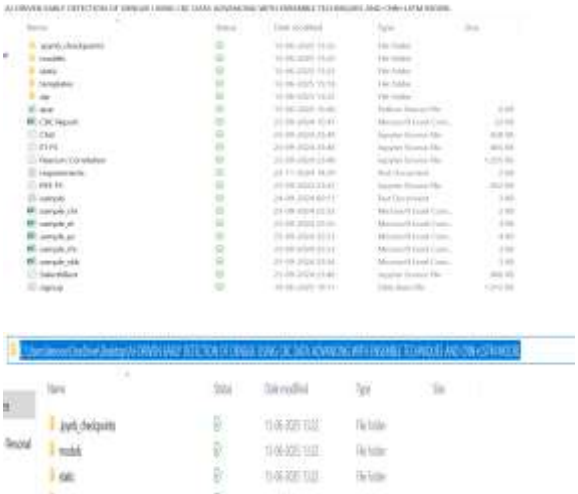


Fig: home page

Fig: Logout page

EXECUTION PROCEDURE



Copy the Folder path

select Anaconda prompt (Anaconda3) and open the app

Open the anaconda prompt

Paste the folder path

Copy the URL

Paste the URL in browser

Dashboard userInterface:

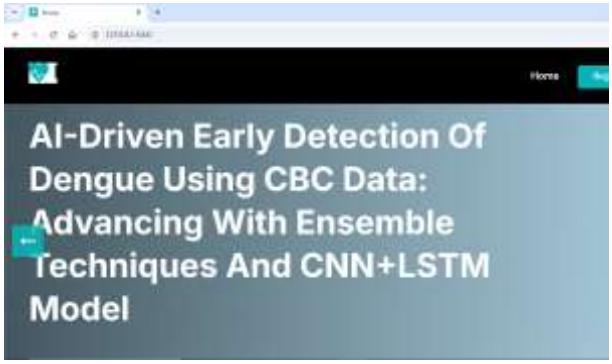
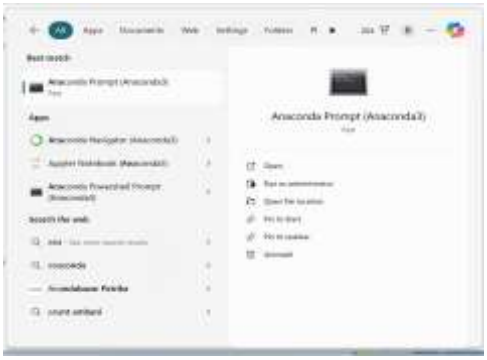


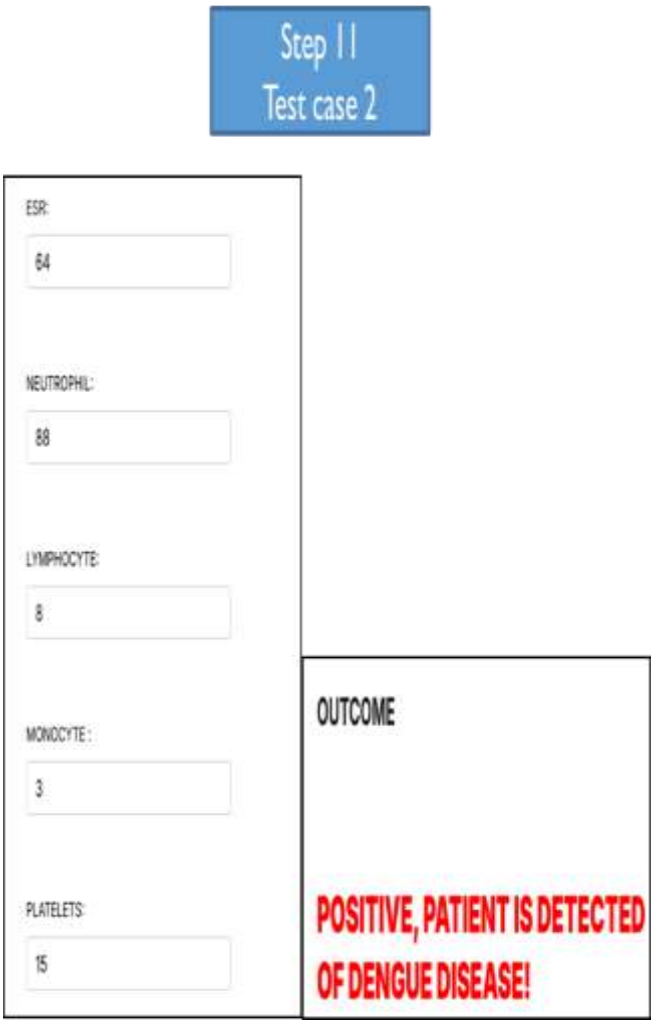
***Fig*: Home page**

Enter the details of user



Give login details like username and password and click login button

After registration click prediction option and select the option PEARSON Coff




Prediction
Graph
Notebook
Logout
Welcome To Dashboard
Pearson Coff
Chi2
RFE FS
SelectKBest
ExtraTree FS
AI-Driven Early Detect
Dengue Using CBC Data:


Step 13
Test case 2
GENDER:
0
HAEMOGLOBIN:
8.6
WBC:
26.9
NEUTROPHIL:
88
MONOCYTE:
3
EOSINOPHIL:
1
RBC:
2.78
PLATELETS:
15
Predict
OUTCOME
POSITIVE, PATIENT IS DETECTED OF DENGUE DISEASE!


Step 13
Test case 1
FORM
GENDER:
0
HAEMOGLOBIN:
12.24948771
WBC:
8.063267591
NEUTROPHIL:
88
MONOCYTE:
3.987286545
EOSINOPHIL:
3.812713455
RBC:
4.576151871
PLATELETS:
278
Predict
OUTCOME
NEGATIVE, PATIENT IS NOT SUFFERING FROM DENGUE DISEASE!

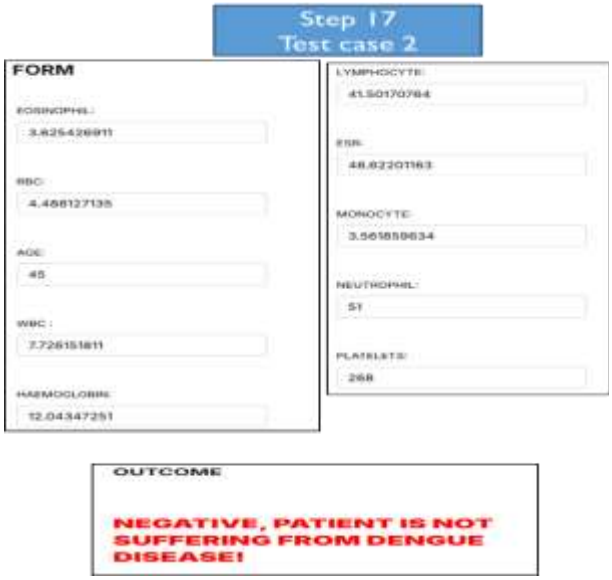

Prediction
Graph
Notebook
Logout
Welcome To Dashboard
Pearson Coff
Chi2
RFE FS
SelectKBest
ExtraTree FS
AI-Driven Early Detect
Dengue Using CBC Data:


Step 15
Test case 1
FORM
HAEMOGLOBIN:
8.6
ESR:
64
WBC:
26.9
NEUTROPHIL:
88
LYMPHOCYTE:
8
MONOCYTE:
3
RBC:
2.78
PLATELETS:
15
Predict
OUTCOME
POSITIVE, PATIENT IS DETECTED OF DENGUE DISEASE!

Step 15
Test case 2
FORM
HAEMOGLOBIN:
14.2
ESR:
6
WBC:
8.9
NEUTROPHIL:
66
LYMPHOCYTE:
27
MONOCYTE:
4
RBC:
4.95
PLATELETS:
280
Predict
OUTCOME
NEGATIVE, PATIENT IS NOT SUFFERING FROM DENGUE DISEASE!


Prediction
Graph
Notebook
Logout
Welcome To Dashboard
AI-Driven Early Detect


Step 17
Test case 1
FORM
EOSINOPHIL:
1
RBC:
2.78
AGE:
22
WBC:
26.9
HAEMOGLOBIN:
8.6
LYMPHOCYTE:
8
ESR:
64
MONOCYTE:
3
NEUTROPHIL:
88
PLATELETS:
15
OUTCOME
POSITIVE, PATIENT IS DETECTED OF DENGUE DISEASE!



CONCLUSION

In conclusion, this artificial intelligence-based system for early detection of dengue using CBC data demonstrates significant potential in improving healthcare outcomes through timely and accurate diagnosis. By leveraging a diverse array of machine learning and deep learning algorithms, including Logistic Regression, SVM, Naive Bayes, Random Forest, AdaBoost, XGBoost, MLP, LightGBM, and advanced ensemble techniques like the Stacking and Voting Classifiers, the system achieves robust predictions. The implementation of deep learning models like CNN, Bi-LSTM, and Transformer further enhances the model's capability to capture complex patterns within the data. SMOTE sampling addresses class imbalances, while feature selection using ExtraTree optimizes model performance. With an accuracy rate of up to 98% using the Voting Classifier, the system proves highly reliable for dengue detection. Additionally, a user-friendly frontend built with Flask ensures ease of use, while secure user authentication guarantees data privacy. This system paves the way for more effective dengue detection and real-time diagnostics, potentially reducing the spread of the disease and improving patient outcomes through early intervention. Further exploration of ensemble techniques and model refinement can continue to enhance the system's capabilities.

7.2 FUTURE ENCHANCEMENT:

The future enchancement of this system includes integrating real-time data from health monitoring devices to enhance early detection capabilities. Additionally, expanding the model to incorporate more diverse datasets can improve its generalizability across different populations. Exploring advanced ensemble techniques and hybrid models may further boost accuracy and robustness. Implementing cloud-based solutions can facilitate scalability and accessibility for healthcare providers. Finally, continuous updates and refinements of the models based on new research findings will ensure the system remains at the forefront of dengue detection technology.

REFERENCES

[1] J. K. Chaw, S. H. Chaw, C. H. Quah, S. Sahrani, M. C. Ang, Y. Zhao, and T. T. Ting, "A predictive analytics model using machine learning algorithms to estimate the risk of shock development among dengue patients," Healthcare Anal., vol. 5, pp. 1–17, Jul. 2024.

[2] T. Akter, M. T. Islam, M. F. Hossain, and M. S. Ullah, "A comparative study between time series and machine learning technique to predict dengue fever in Dhaka city," Discrete Dyn. Nature Soc., vol. 2024, pp. 1–12, May 2024.

[3] S. Roy, A. Biswas, M. T. A. Shawon, S. Akter, and M. M. Rahman, ‘‘Land use and meteorological influences on dengue transmission dynamics in Dhaka city, Bangladesh,’’ Bull. Nat. Res. Centre, vol. 48, no. 1, pp. 1–24, Mar. 2024, doi: 10.1186/S42269-024-01188-0.

[4] N. Ali, ‘‘The recent burden of dengue infection in bangladesh: A serious public health issue,’’ J. Infection Public Health, vol. 17, no. 2, pp. 226–228, Feb. 2024.

[5] N. Sharif, N. Sharif, A. Khan, and S. K. Dey, ‘‘The epidemiologic and clinical characteristics of the 2023 dengue outbreak in Bangladesh,’’ Open Forum Infectious Diseases, vol. 11, no. 2, pp. 1–29, Feb. 2024, doi: 10.1093/OFID/OFAE066.

[6] D. C. Kajeguka, F. M. Mponela, E. Mkumbo, A. N. Kaaya, D. Lasway, R. D. Kaaya, M. Alifrangis, E. Elanga-Ndille, B. T. Mmbaga, and R. Kavishe, ‘‘Prevalence and associated factors of dengue virus circulation in the rural community, Handeni district in Tanga, Tanzania,’’ J. Tropical Med., vol. 2023, pp. 1–9, Nov. 2023.

[7] M. A. Kabir, H. Zilouchian, M. A. Younas, and W. Asghar, ‘‘Dengue detection: Advances in diagnostic tools from conventional technology to point of care,’’ Biosensors, vol. 11, no. 7, p. 206, Jun. 2021.

[8] C. Davi, A. Pastor, T. Oliveira, F. B. d. L. Neto, U. Braga-Neto, A. W. Bigham, M. Bamshad, E. T. A. Marques, and B. Acioli-Santos, ‘‘Severe dengue prognosis using human genome data and machine learning,’’ IEEE Trans. Biomed. Eng., vol. 66, no. 10, pp. 2861–2868, Oct. 2019.

[9] D. Sarma, S. Hossain, T. Mittra, Md. A. M. Bhuiya, I. Saha, and R. Chakma, ‘‘Dengue prediction using machine learning algorithms,’’ in Proc. IEEE 8th R10 Humanitarian Technol. Conf., Dec. 2020, pp. 1–6.

[10] E. Fernández, M. Smieja, S. D. Walter, and M. Loeb, ‘‘A predictive model to differentiate dengue from other febrile illness,’’ BMC Infectious Diseases, vol. 16, no. 1, pp. 1–7, Dec. 2016.

[11] H. Mayrose, G. M. Bairy, N. Sampathila, S. Belurkar, and K. Saravu, ‘‘Machine learning-based detection of dengue from blood smear images utilizing platelet and lymphocyte characteristics,’’ Diagnostics, vol. 13, no. 2, p. 220, Jan. 2023.

[12] S. Sabrina Prome, T. Basak, T. Islam Plabon, and R. Khan, ‘‘Prediction of dengue cases in Bangladesh using explainable machine learning approach,’’ in Proc. Int. Conf. Inventive Comput. Technol. (ICICT), Apr. 2024, pp. 1–5.

[13] J. D. Mello-Román, J. C. Mello-Román, S. Gómez-Guerrero, and M. García-Torres, ‘‘Predictive models for the medical diagnosis of dengue: A case study in Paraguay,’’ Comput. Math. Methods Med., vol. 2019, pp. 1–7, Jul. 2019.

[14] S. K. Dey, M. M. Rahman, A. Howlader, U. R. Siddiqi, K. M. M. Uddin, R. Borhan, and E. U. Rahman, ‘‘Prediction of dengue incidents using hospitalized patients, metrological and socio-economic data in Bangladesh: A machine learning approach,’’ PLoS One, vol. 17, no. 7, Jul. 2022, Art. no. e0270933.

[15] B. Abdualgalil, S. Abraham, and W. M. Ismael, ‘‘Early diagnosis for dengue disease prediction using efficient machine learning techniques based on clinical data,’’ J. Robot. Control (JRC), vol. 3, no. 3, pp. 257–268, May 2022.

Authors:

Mr. B. Amarnath Reddy is an Assistant Professor in the Department of Master of Computer Applications at QIS College of Engineering and Technology, Ongole, Andhra Pradesh. He earned his M.Tech from Vellore Institute of Technology(VIT), Vellore. His research interests include Machine Learning,Programming Languages. He is committed to advancing research and fostering innovation while mentoring students to excel in both academic and professional pursuits.