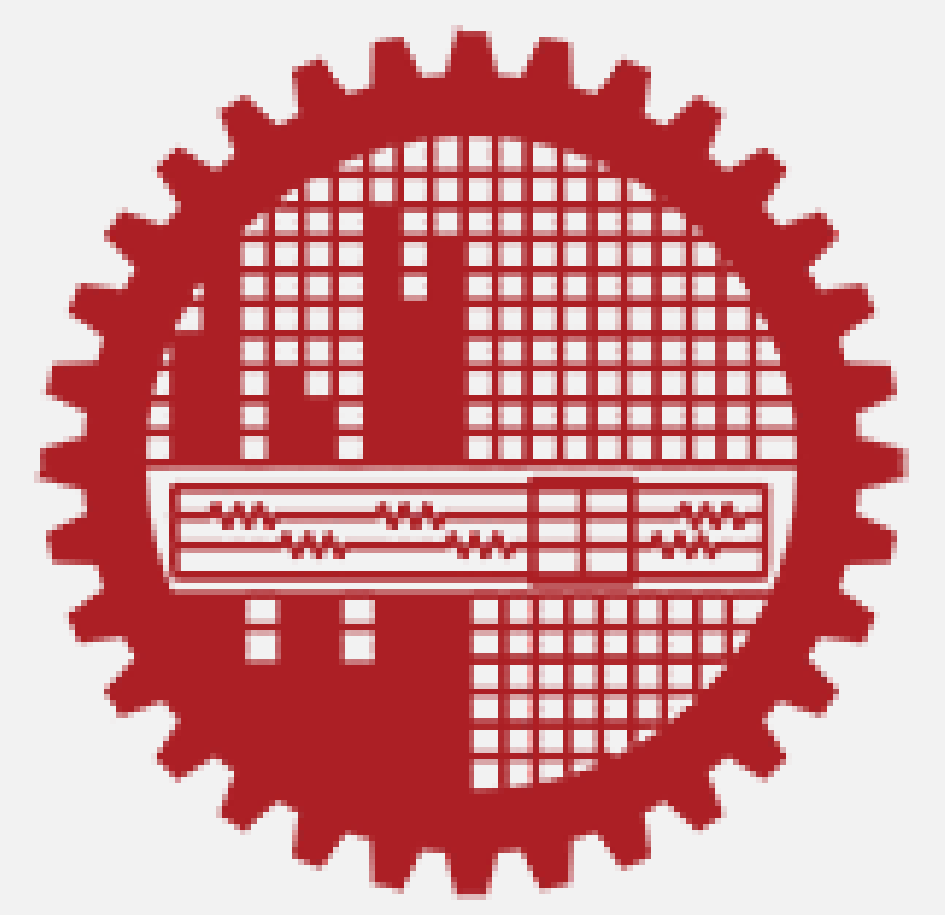


DEVELOPMENT OF A LOW-COST DATA LOGGER FOR TRIAXIAL TEST OF SOIL

Md. Masudur Rahman, Dr. Md. Liakot Ali



Abstract

This project endeavors to tackle the challenges associated with data logging in soil mechanics laboratories by proposing the development of a microcontroller-based, low-cost data logger. The primary goal is to establish an affordable alternative to commercial systems, facilitating automated monitoring of soil parameters during experiments such as consolidation, uniaxial compression, triaxial compression, and shear box tests.

The methodology involves integrating a suitable microcontroller with sensors for measuring relevant soil parameters, along with signal conditioning to ensure accuracy. Non-volatile memory is employed for extended data storage, and a communication module enables real-time data transfer. Efficient power management, noise cancellation algorithms, and a robust calibration system contribute to the reliability of the data logger.

The user-friendly interface simplifies configuration, while extensive testing under various soil mechanics scenarios validates the system's performance. Comprehensive documentation is provided, and an open-source approach is considered to encourage collaboration and further development within the scientific and engineering communities. This project aims to democratize access to automated data logging systems in soil mechanics, fostering efficiency and affordability in academic and research settings..

Background & Motivation

Data loggers are electronic devices crafted for the measurement and recording of physical or electrical parameters over time, employing external instruments and sensors. They serve diverse purposes in scientific and engineering domains, spanning applications in in-vehicle data logging, environmental monitoring, structural health monitoring, and machine condition monitoring.

In civil engineering soil mechanics labs, accurate sensor readings are vital for tests like consolidation and compression. Continuous monitoring of soil response is crucial for comprehensive data. However, manual recording by technicians is cumbersome and error-prone, risking productivity losses and inaccuracies.

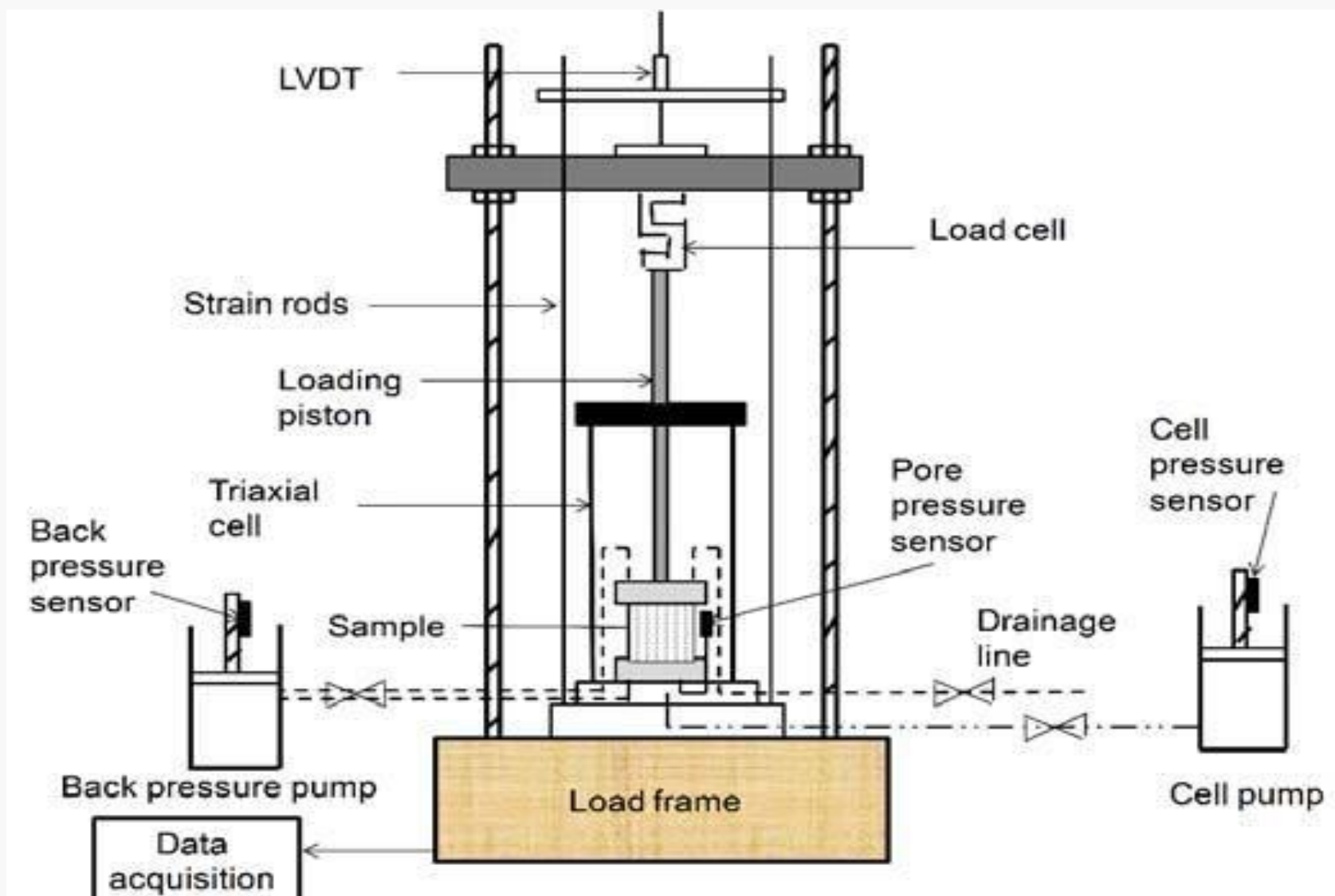


Figure-1: Schematic representation of a Triaxial Test apparatus.

Recognizing challenges in manual recording and expensive automated systems, the project aims to create a low-cost, microcontroller-based data logger for soil mechanics. This initiative is fueled by the need for an accessible solution tailored to soil mechanics experiments. The local development aims to enhance efficiency and widen access to data loggers in educational and research institutions by overcoming financial constraints associated with commercial solutions.

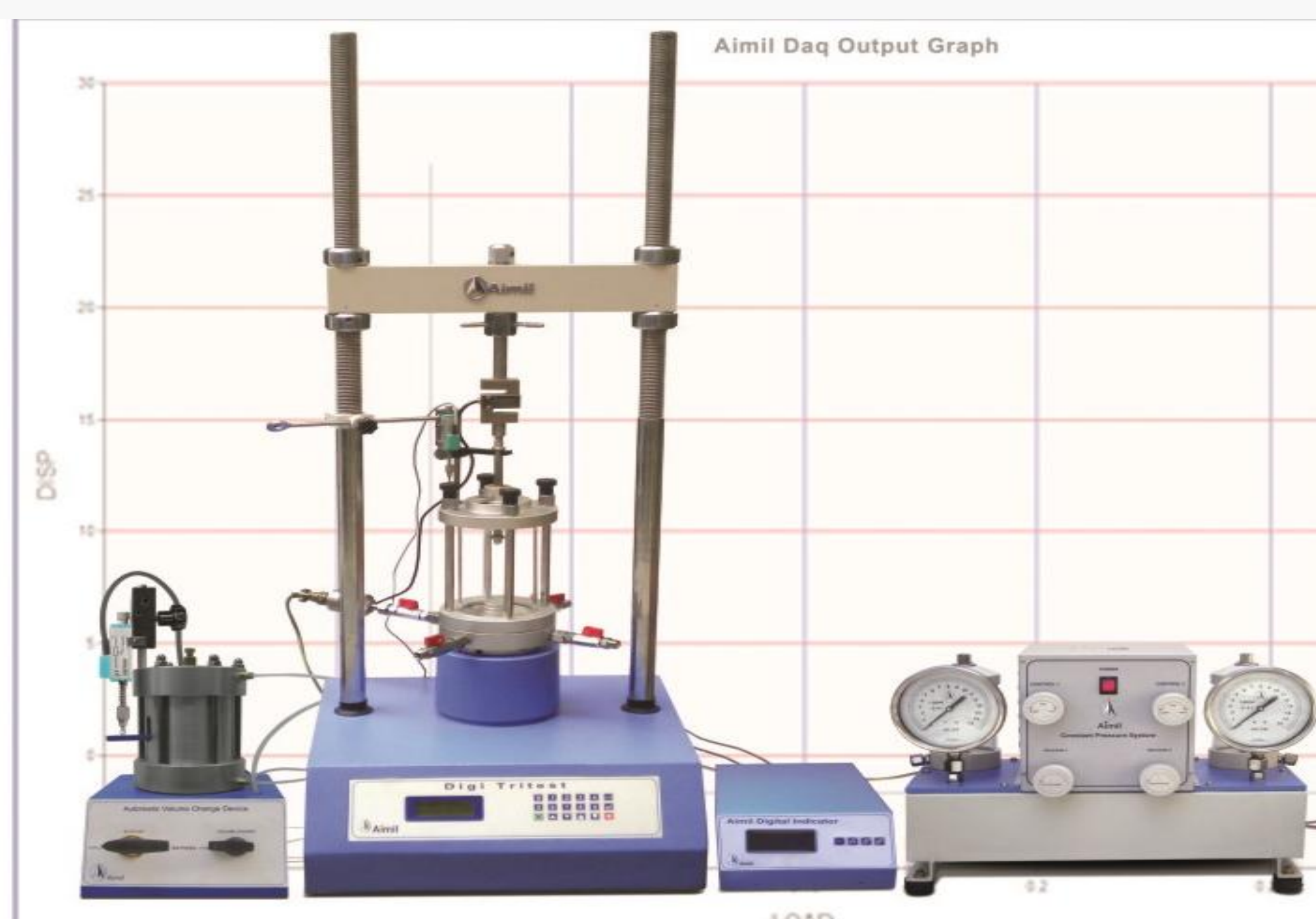


Figure-2: Instrumentation for Triaxial Test apparatus (Ref: Aimil)

Proposed Idea and Methodology

Idea: Addressing the limitations of expensive and inefficient manual data recording in soil mechanics laboratories, this project proposes a novel, low-cost data logger built around a microcontroller. This affordable alternative seamlessly integrates with common experiments like triaxial compression, consolidation, and direct shear tests, continuously recording and storing sensor readings throughout extended durations. By automating data acquisition, the logger eliminates human error while ensuring accurate and reliable measurements of vital soil parameters like stress and deformation. This not only streamlines research workflows but also paves the way for more complex and insightful experiments, ultimately fostering advancements in geotechnical engineering through improved data accuracy, reduced labor costs, and deeper understanding of soil behavior under load.

Methodology:

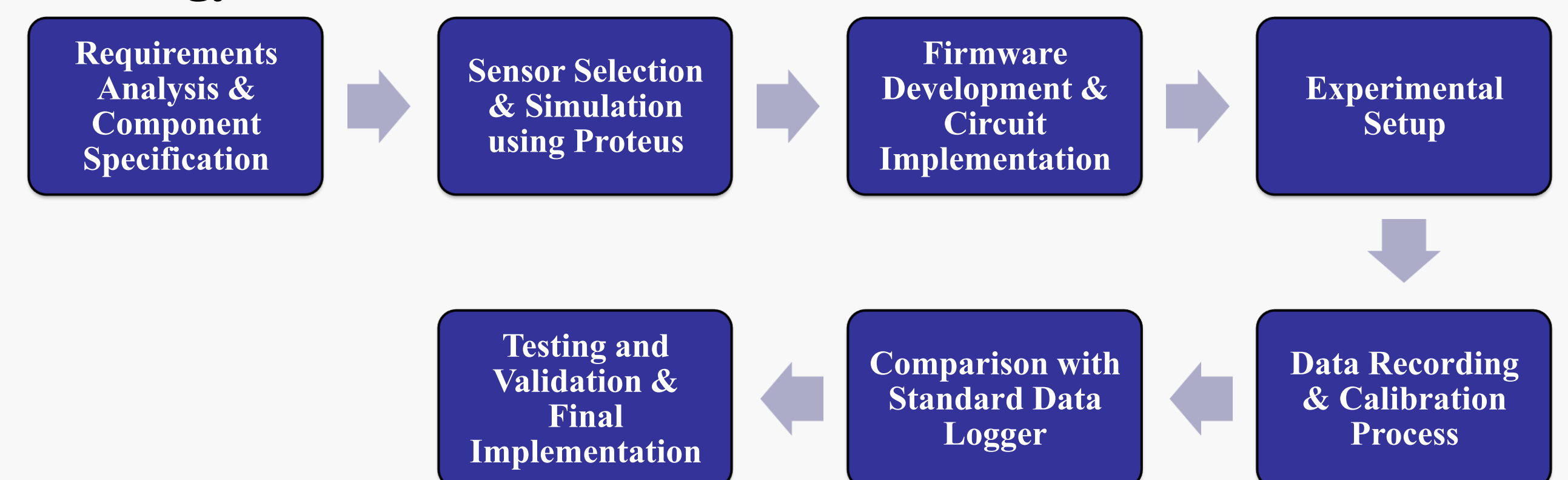


Figure-3: Flow chart of methodology

The methodology begins with a meticulous analysis of project requirements, with subsequent finalization of hardware and software specifications. High-precision sensors, LVDT and S-Type Load Cell 10 kN, are chosen for data acquisition. Circuit design proceeds via Proteus software simulation, followed by firmware development for the AVR microcontroller. The breadboard prototype is constructed, replicating soil test conditions for experimental evaluation.

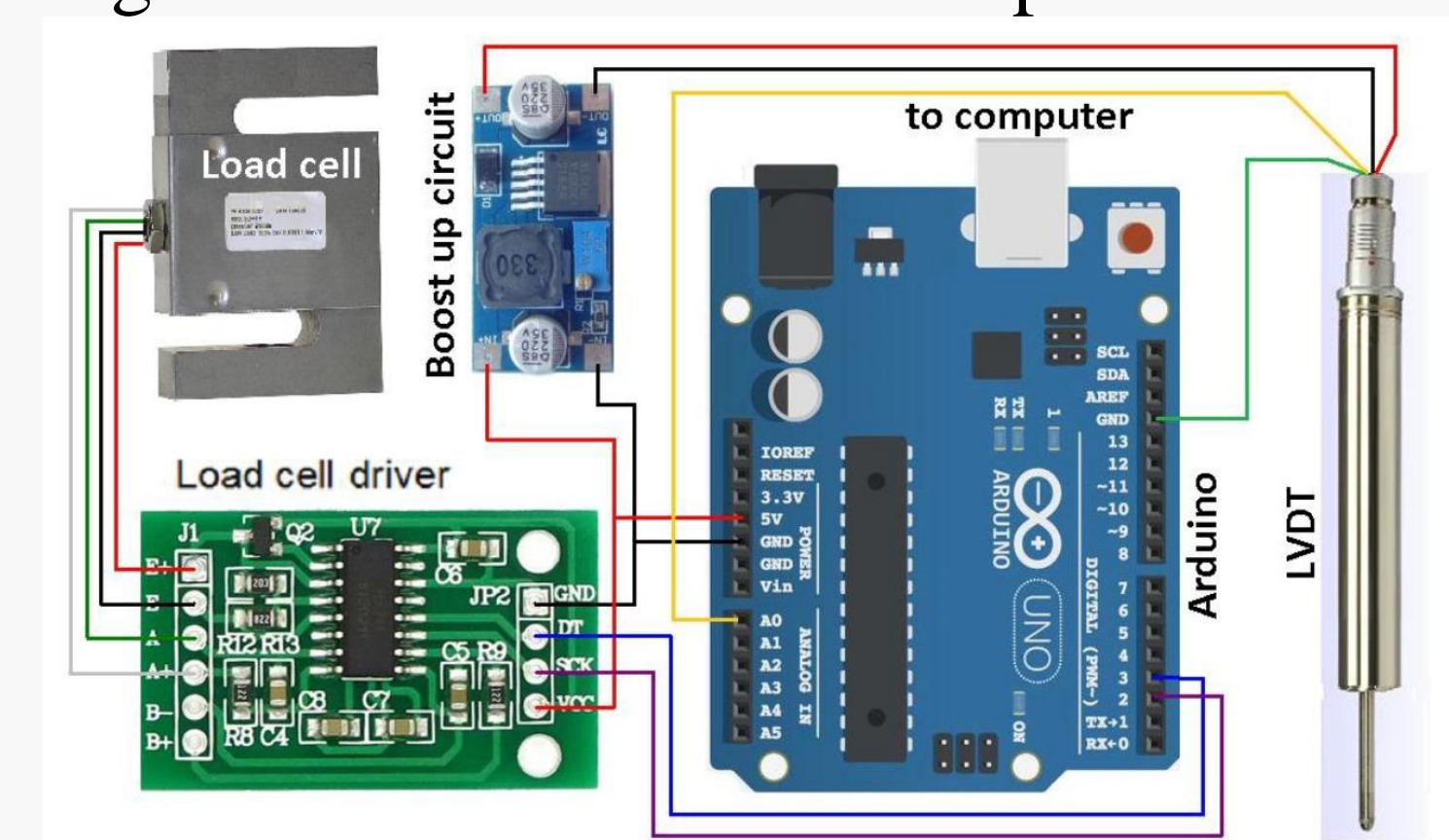


Figure-4: Design of Circuit

An SD card module is integrated for data recording, and calibration is performed against reference instruments. Comprehensive testing demonstrates reliable performance under various scenarios. Thorough documentation covers specifications, design details, code, and calibration procedures. Continuous improvement based on testing feedback leads to the final implementation, ready for integration into soil mechanics laboratories..

Results

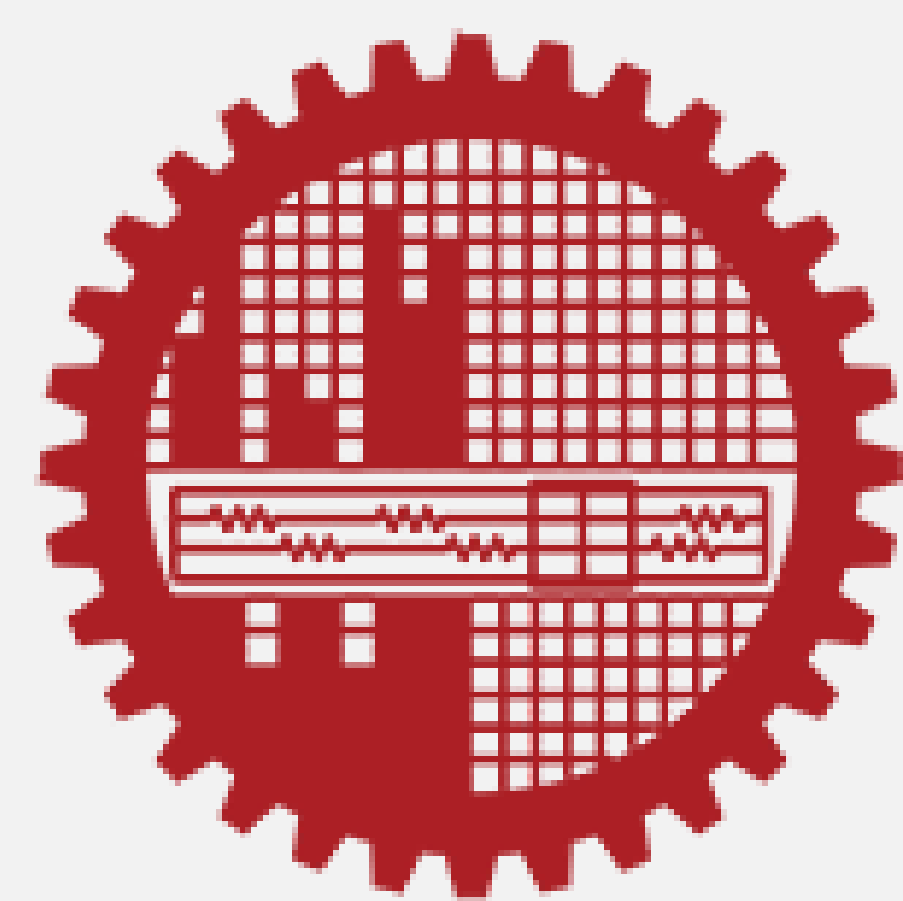
- The microcontroller-based data logger designed for soil mechanics experiments demonstrated reliable and better performance during rigorous testing.
- Integrated with recommended sensors, the system provided precise measurements in consolidation, compression, and shear box tests.
- The user-friendly interface and SD card storage further enhance its effectiveness.
- Overall, the results highlight the success of this cost-effective solution in addressing challenges associated with manual recording in soil mechanics laboratories.

Future work

- Integration of additional sensors to expand the range of measured parameters.
- Further enhancements in communication modules can be investigated for increased flexibility.
- Advanced calibration techniques may be researched and implemented to refine the accuracy of displacement and load measurements.

A New Perspective on Assisting the Visually Impaired with Navigation and Awareness

Sheezanul Hassan, Md. Liakot Ali, Muhammad Sheikh Sadi



Abstract

Safe and independent mobility is one of the major daily challenges faced by the visually impaired. They struggle with object detection and obstacle avoidance, making it challenging for them to perform their day-to-day tasks. However, establishing secure and safe mobility for the visually impaired is a difficult undertaking that must be done accurately and efficiently. Recognizing currency is another severe problem that the visually impaired confront because different notes, in our country, have similar surfaces and sizes. In this research, we propose a currency and object detection and relative position identification system that will benefit the visually impaired to detect objects and Bangladeshi currency notes and make the appropriate decision regarding their situation in both indoor and outdoor aspects. The model proposed in the system has been evaluated with real-world objects for evaluating the performance of the proposed method. The experimental analysis shows that the system has an average accuracy of 99.44%, a recall of 99.38%, a precision of 99.39%, and an F1 score of 99.37% after experimenting with 5281 images of real objects.

Background & Motivation

- ❖ Around 43 million people are blind and around 295 million people in the world have moderate to severe vision impairment
- ❖ Visually impaired people fail to detect and avoid obstacles in their path, thus causing them emotional suffering, undercutting their independence.
- ❖ Identifying currency-note is another major issue for the visually impaired.
- ❖ Smart Cane, Ultra-Cane, Ray Mobility Aid are some of the commercially available devices for visually impaired people.
- ❖ The acceptance rate of these devices is relatively low due to the high cost, inaccuracy for detecting obstacle types, the identification of different objects in both daylight and dark environments, and the detection of currency banknotes
- ❖ There is room for research to create a stand-alone system that can handle the objects and currency detection system for the visually impaired.

Proposed Methodology

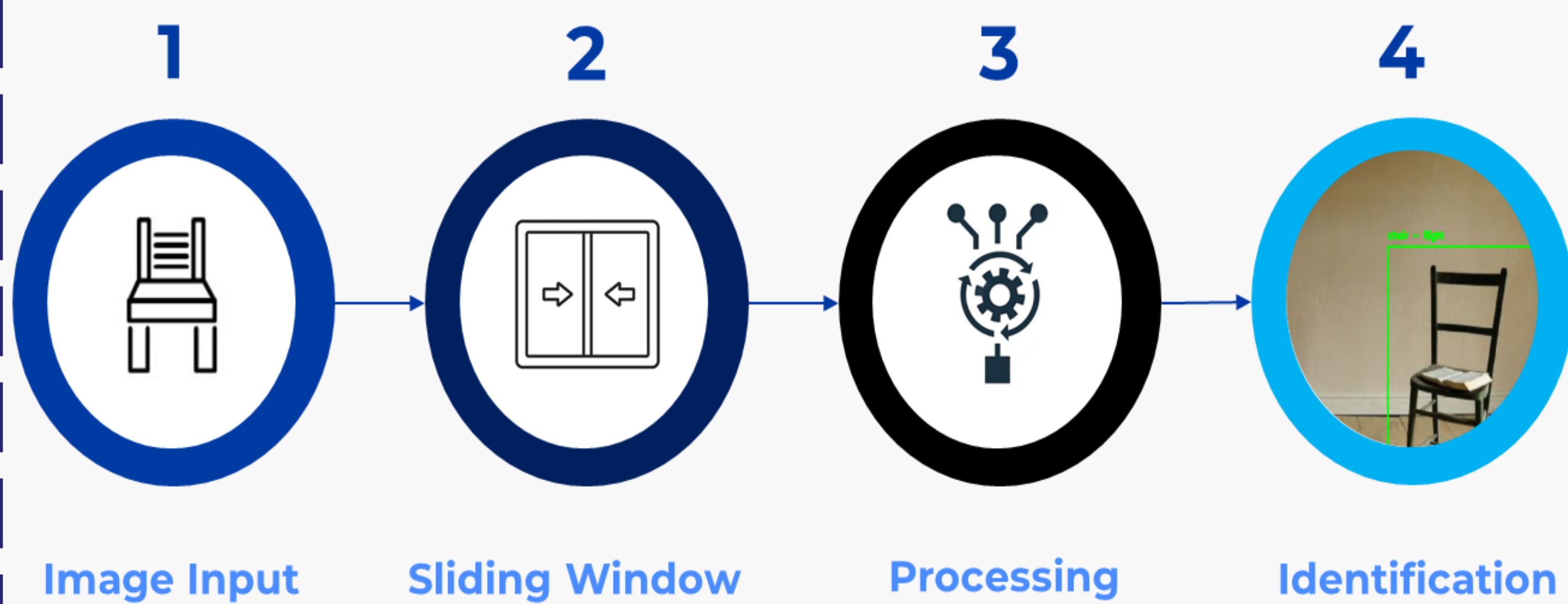


Figure 2. Relative position and object identification

Proposed Methodology

The proposed method is summarized in Figure 1

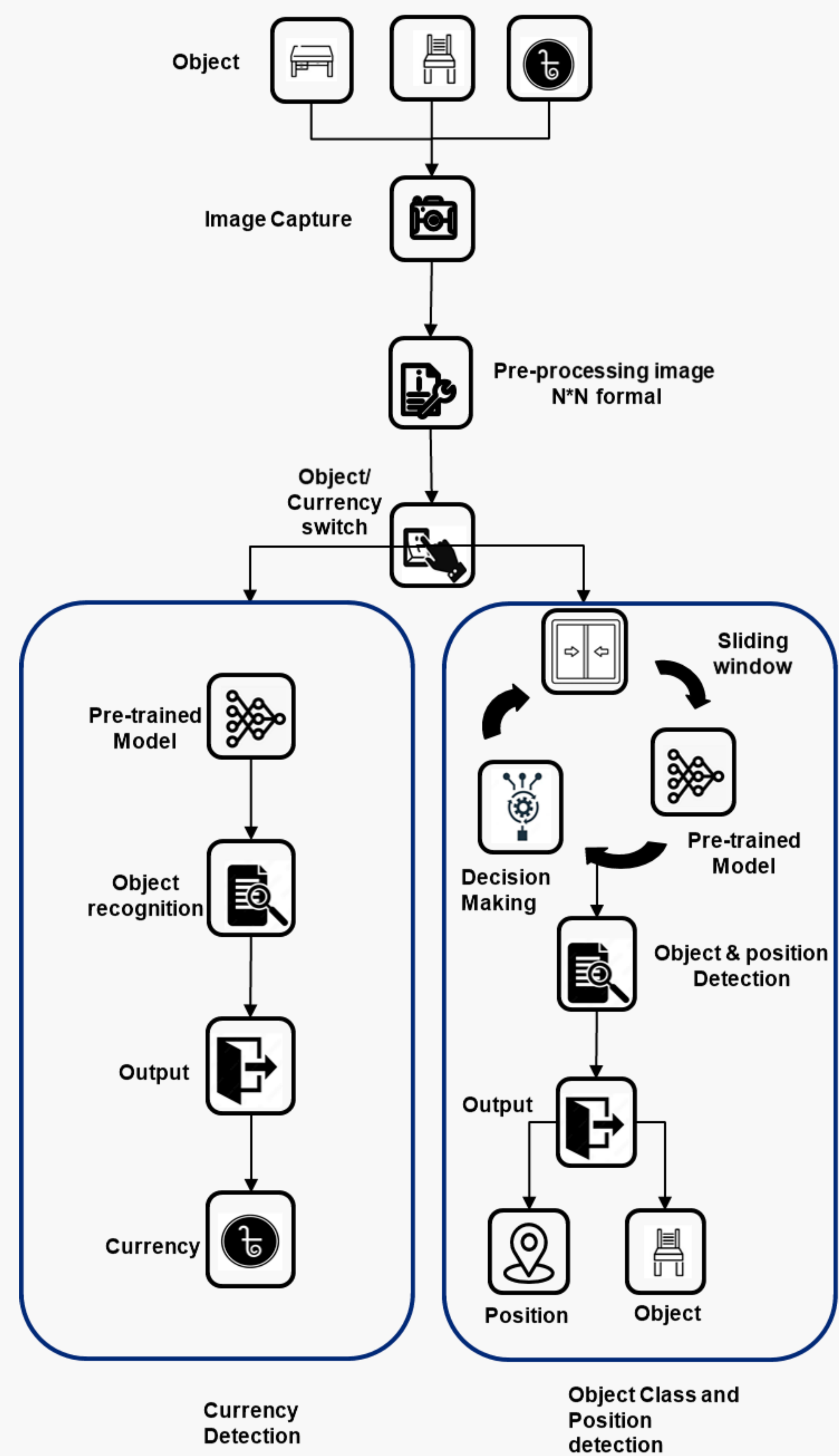


Figure 1. The Proposed CNN Based System Model Diagram

Results

- ❖ Weighted average, precision, recall, and F1 score for all five classes for object detection are 99.826%, 99.825%, and 99.825% respectively on 1144 samples. Confusion matrix is shown in Figure 4.
- ❖ Overall values of precision, recall, F1 score, and accuracy of all of the fifteen classes for currency detection are 99.39%, 99.38%, 99.37%, and 99.44% respectively. Confusion matrix is shown in Figure 5

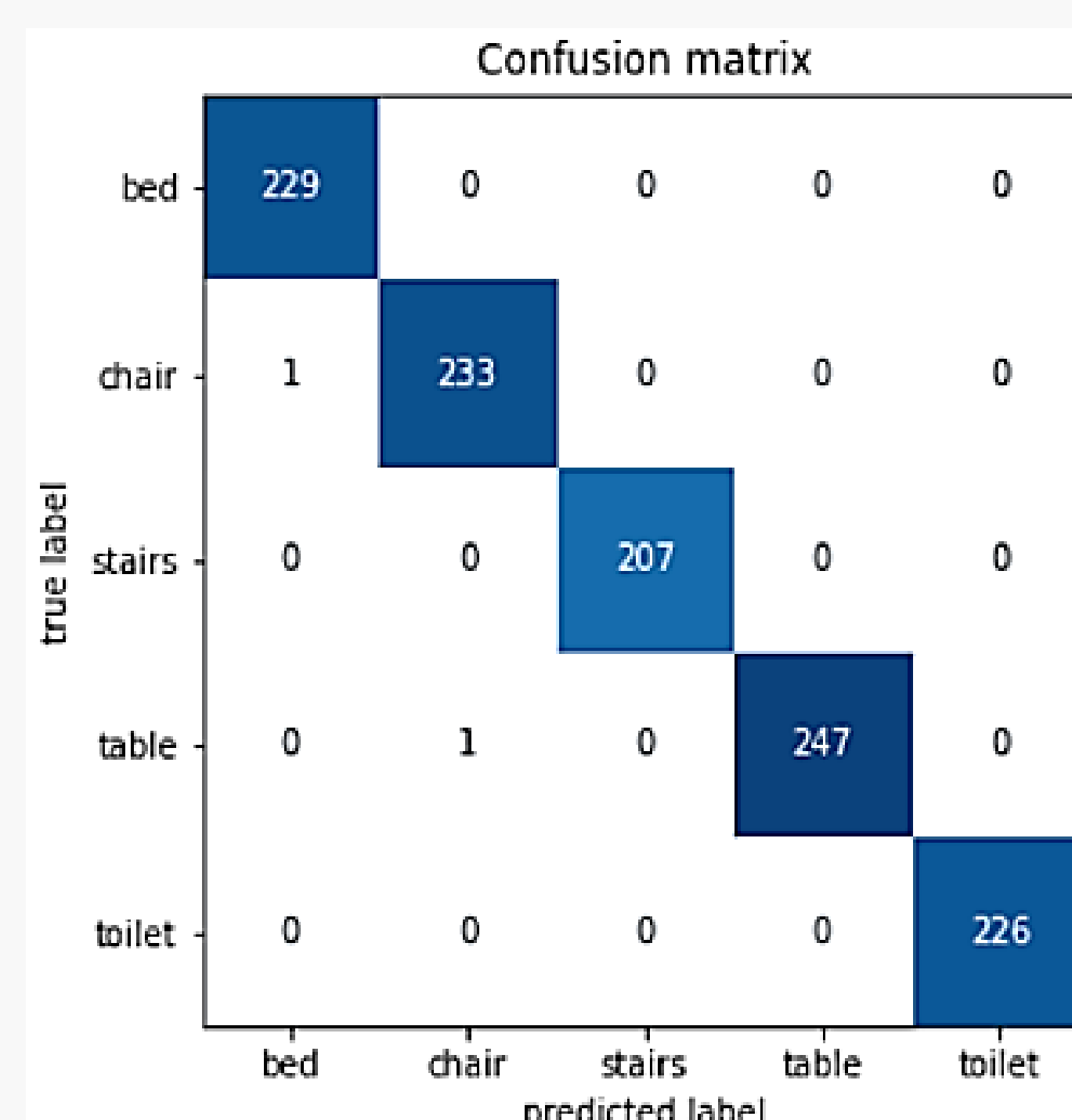


Figure 3. Confusion matrix for object detection

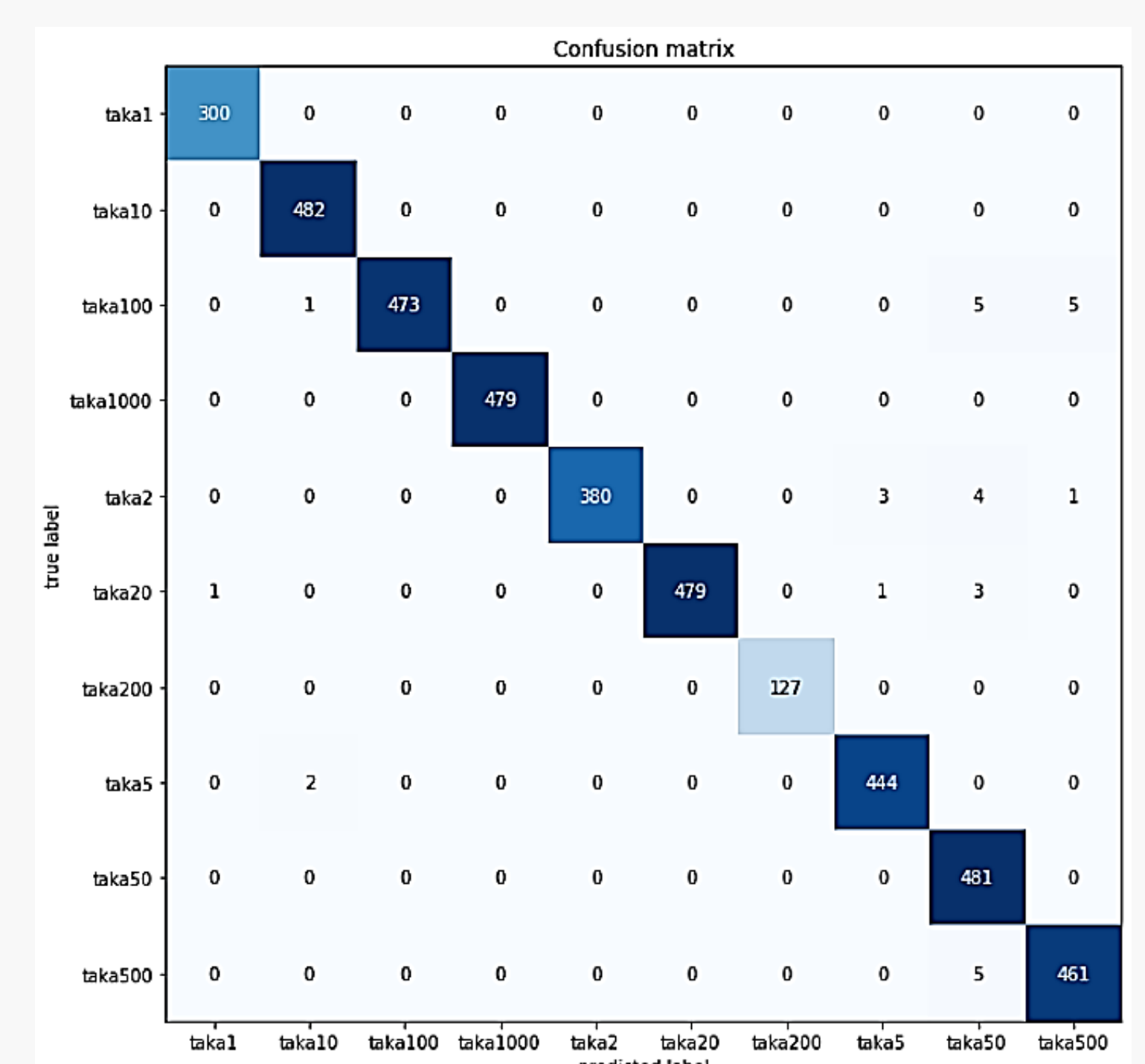
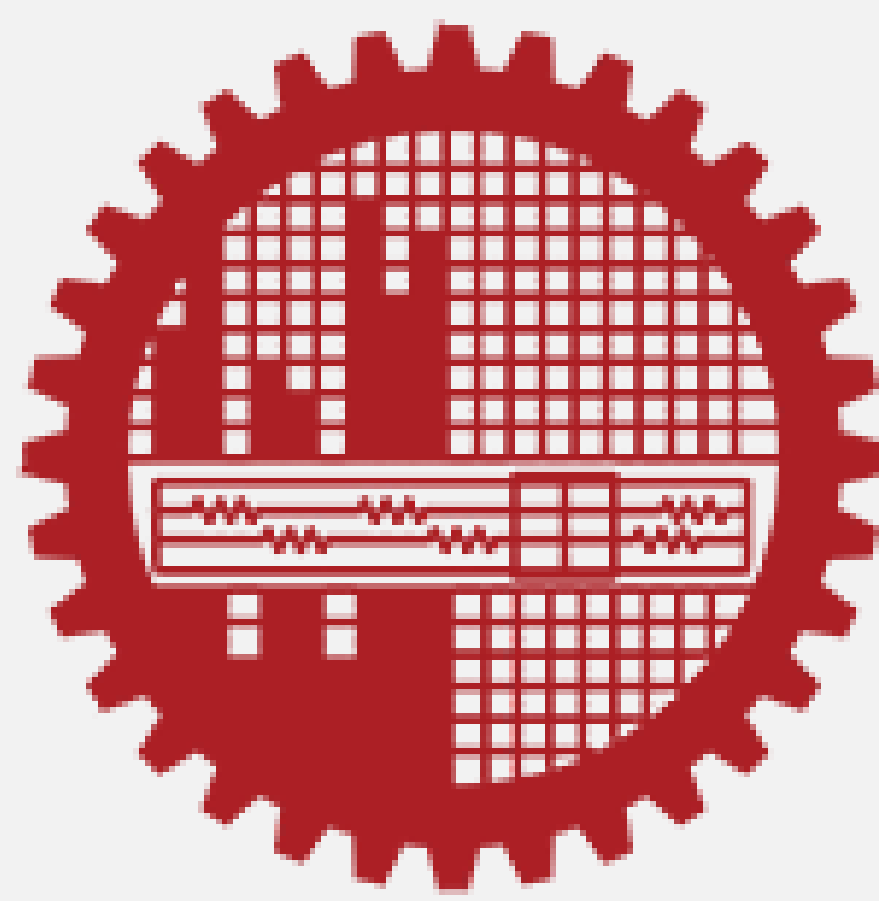


Figure 4. Confusion matrix for currency detection

Developing a Real-time Hand-Gesture Recognition System for Wheelchair Control

Md Rafiul Huda, Md. Liakot Ali, Muhammad Sheikh Sadi



Abstract

People with motor disabilities are increasing day by day in the world. Many disabled and elderly persons (131.8 million currently) need to use wheelchairs to perform regular movements and activities. It becomes challenging to run traditional wheelchairs for those who suffer from dexterity-inhibiting disorders and mobility impairments. Now a days, many smart wheelchair systems have been proposed, but these are not so suitable and flexible for users having dexterity problems. This poster presents the design and development of an efficient hand gesture recognition algorithm for a smart wheelchair control, where a mathematical model is developed utilizing the positions of the hand landmarks and some thresholds. The proposed method has been validated in a real powered wheelchair and the experimental study shows that it outperforms existing methods and it will resolve the issues of other proposed systems.

Background & Motivation

- ❑ The number of wheelchair users is increasing day by day.
- ❑ Safe navigation, independent mobility and low cost are the key issues for wheelchair users, especially those suffering from mobility impairments.
- ❑ Disables people face many problems using traditional wheelchairs which require muscle power and wide movement of hands
- ❑ The advancement of technology has allowed wheelchairs to become intelligent by incorporating contemporary sensors, and AI.
- ❑ Gestures can be easily incorporated into wheelchair functionalities utilizing computer vision and AI.
- ❑ Several researches carried out i.e. joystick, depth camera or android application based, but not suitable and flexible for users having dexterity problems or face difficulties with the movement of hand.
- ❑ Accuracy of existing systems is variable or not so high.
- ❑ Recent researchers used skin segmentation, KCF tracking, CNN model but couldn't show good performance under daylight or if the background matches with skin color.
- ❑ It's still a challenging issue to build an efficient gesture recognition system for wheelchair control that will consider users' flexibility, high accuracy, and will work in any lighting condition of environment.



(a)



(b)

Figure 1. (a) A traditional manual wheelchair *
(b) A Smart powered wheelchair **

* picture from www.google.com

** developed by Sadi, M.S et al.

Proposed Idea and Methodology

- ❑ The proposed methodology comprises three major steps :
 - ❖ Hand Detection and Hand Tracking by a robust framework-MediaPipe
 - ❖ Hand Landmarks Extraction and Distance(normalized) matrix preparation
 - ❖ Gesture Recognition using a Mathematical Model (based on distances & thresholds)

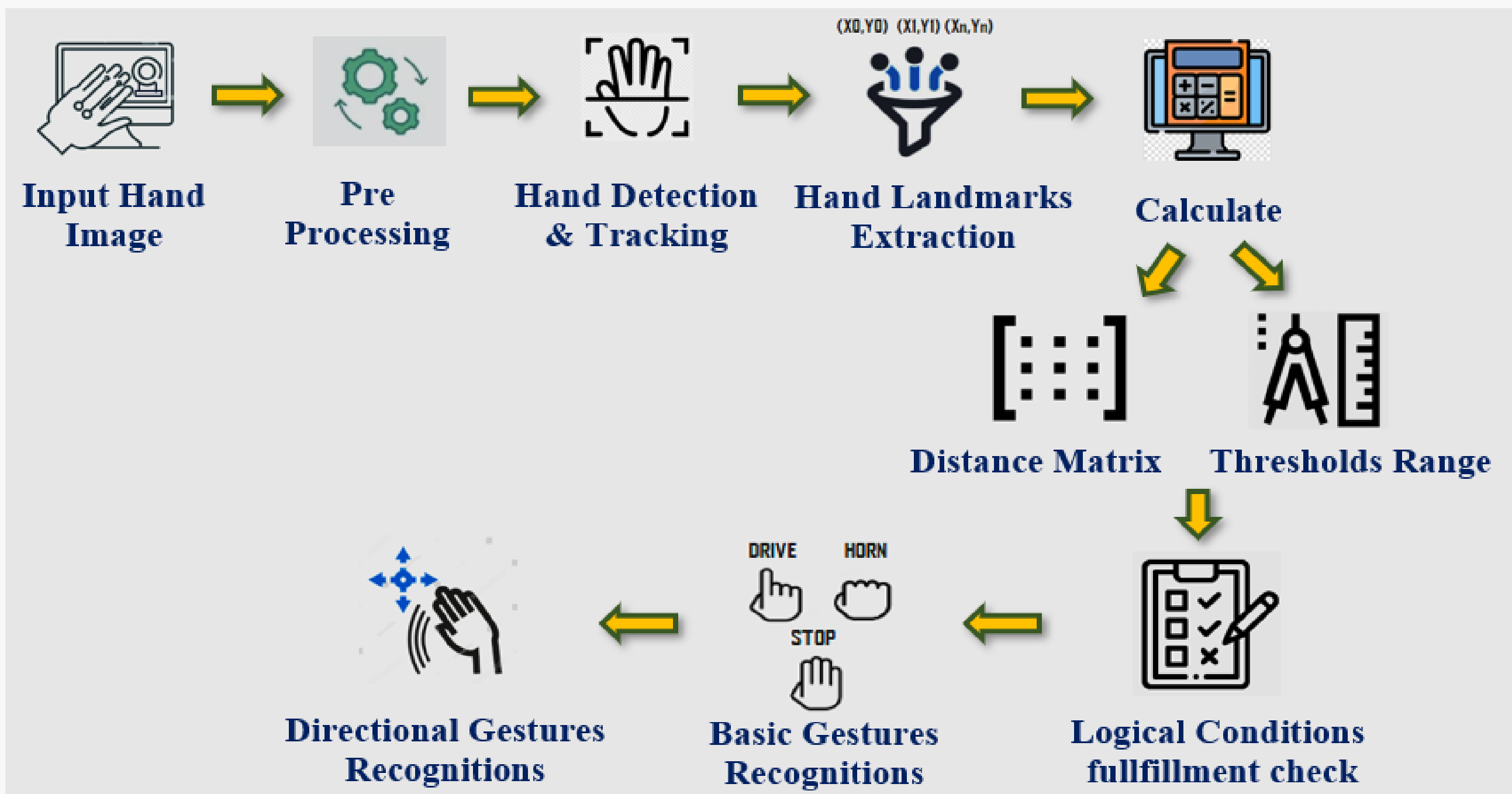


Figure 2. Architectural components of the proposed mathematical mode

- ❑ Three Thresholds are determined from two types of gestures at the beginning for once
- ❑ Different Gestures are recognized upon satisfying the logical conditions with those Thresholds

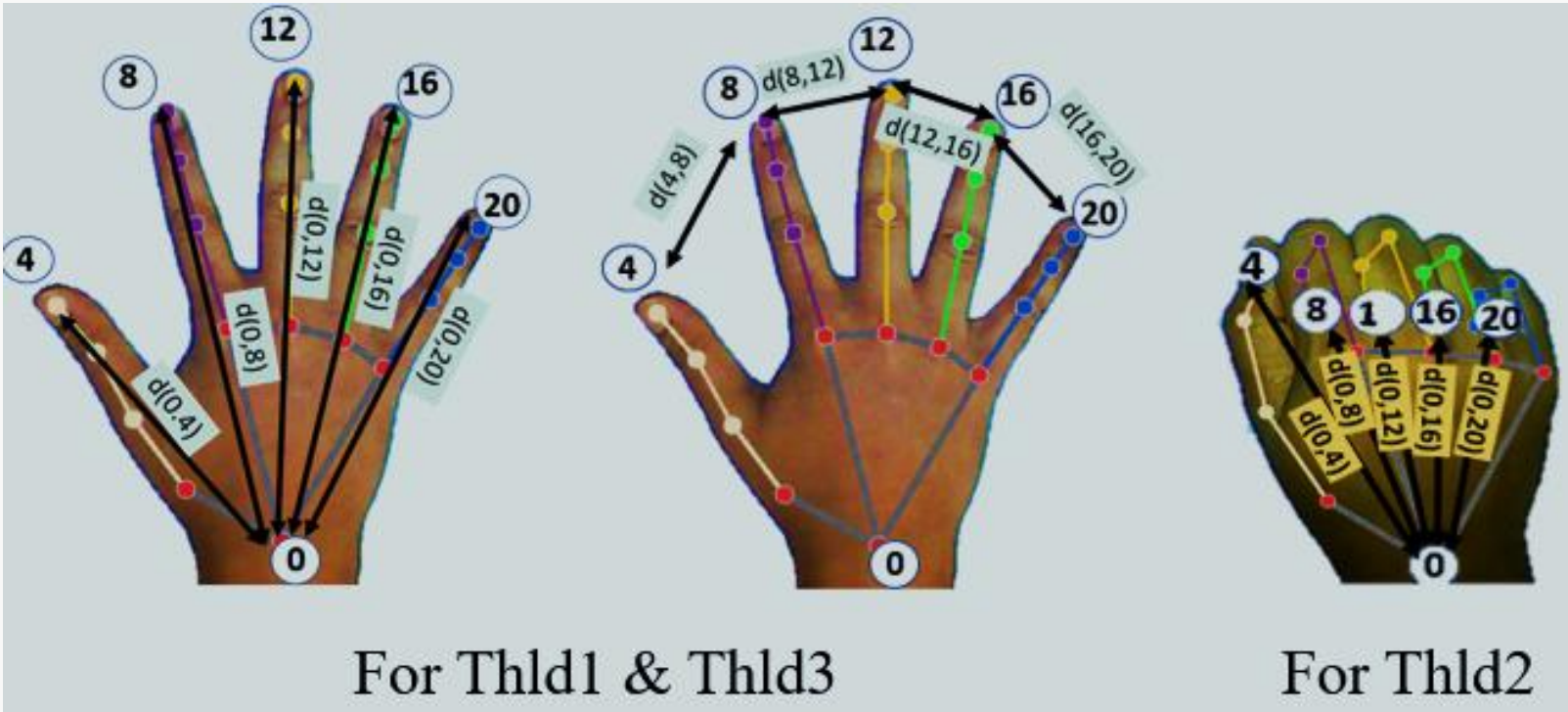


Figure 3. Thresholds determination

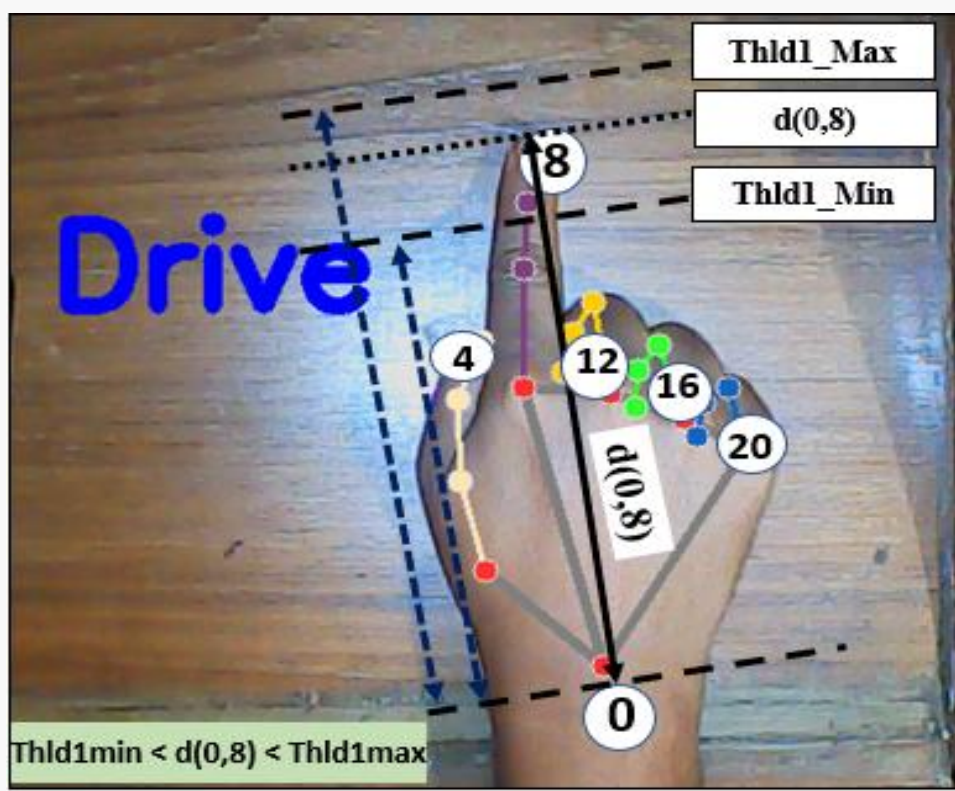


Figure 4. Visualization of hand landmarks and satisfying conditions

Gestures	Distance parameters	Required conditions to meet
Drive	$d(0,8)$	$Thld1_{min} < d(0,8) < Thld1_{max}$
	$d(0,12)$	$Thld2_{min} < d(0,12) < Thld2_{max}$
	$d(0,16)$	$Thld2_{min} < d(0,16) < Thld2_{max}$
	$d(0,20)$	$Thld2_{min} < d(0,20) < Thld2_{max}$
Stop	$d(0,4)$	$d(0,4) < Thld1$
	$d(0,8)$	$d(0,8) > Thld1$
	$d(0,12)$	$d(0,12) > Thld1$
	$d(0,16)$	$d(0,16) > Thld1$
Horn	$d(0,20)$	$d(0,20) > Thld2$
	$d(0,4)$	$d(0,4) > Thld2$
	$d(4,8)$	$Thld3_{min} < d(4,8) < Thld3_{max}$
	$d(8,12)$	$Thld3_{min} < d(8,12) < Thld3_{max}$
Horn	$d(12,16)$	$Thld3_{min} < d(12,16) < Thld3_{max}$
	$d(16,20)$	$Thld3_{min} < d(16,20) < Thld3_{max}$
	$d(0,8)$	$d(0,8) < Thld1_{min}$
	$d(0,12)$	$d(0,12) < Thld2$
	$d(0,16)$	$d(0,16) < Thld2$
	$d(0,20)$	$d(0,20) < Thld2$

Figure 5. Logical conditions to be met for basic gestures recognition

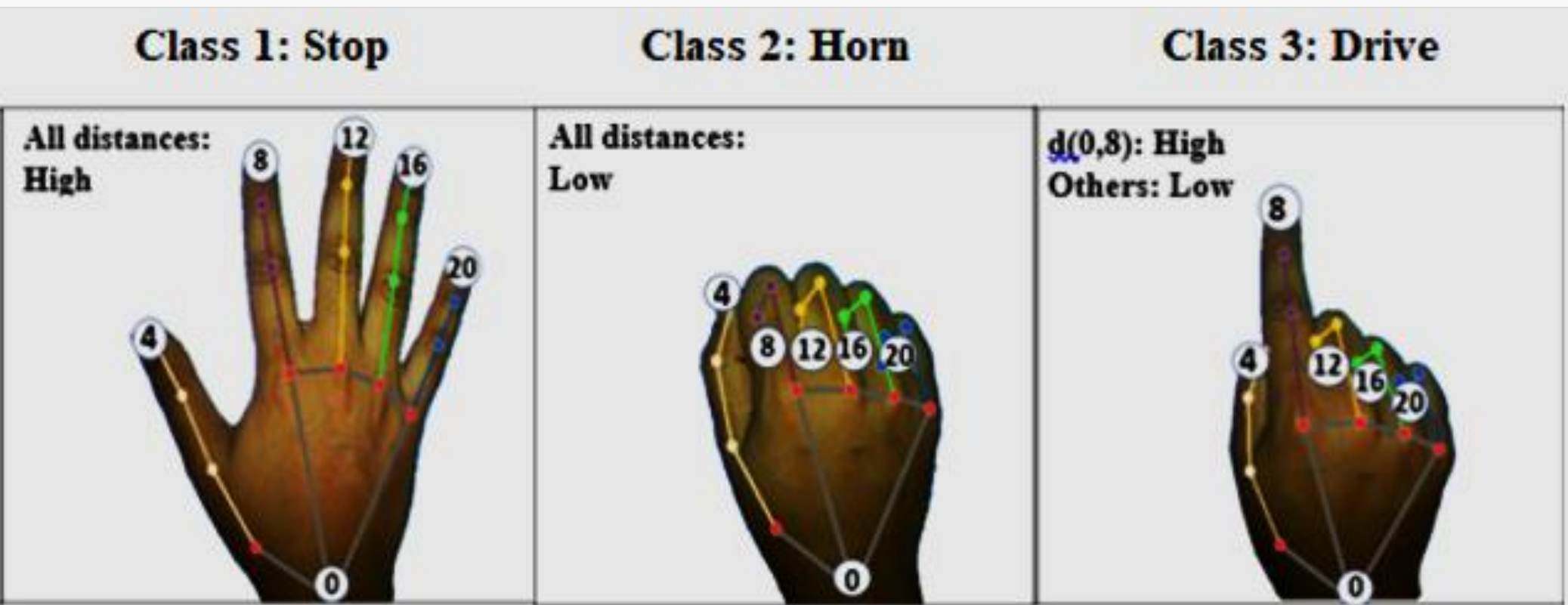
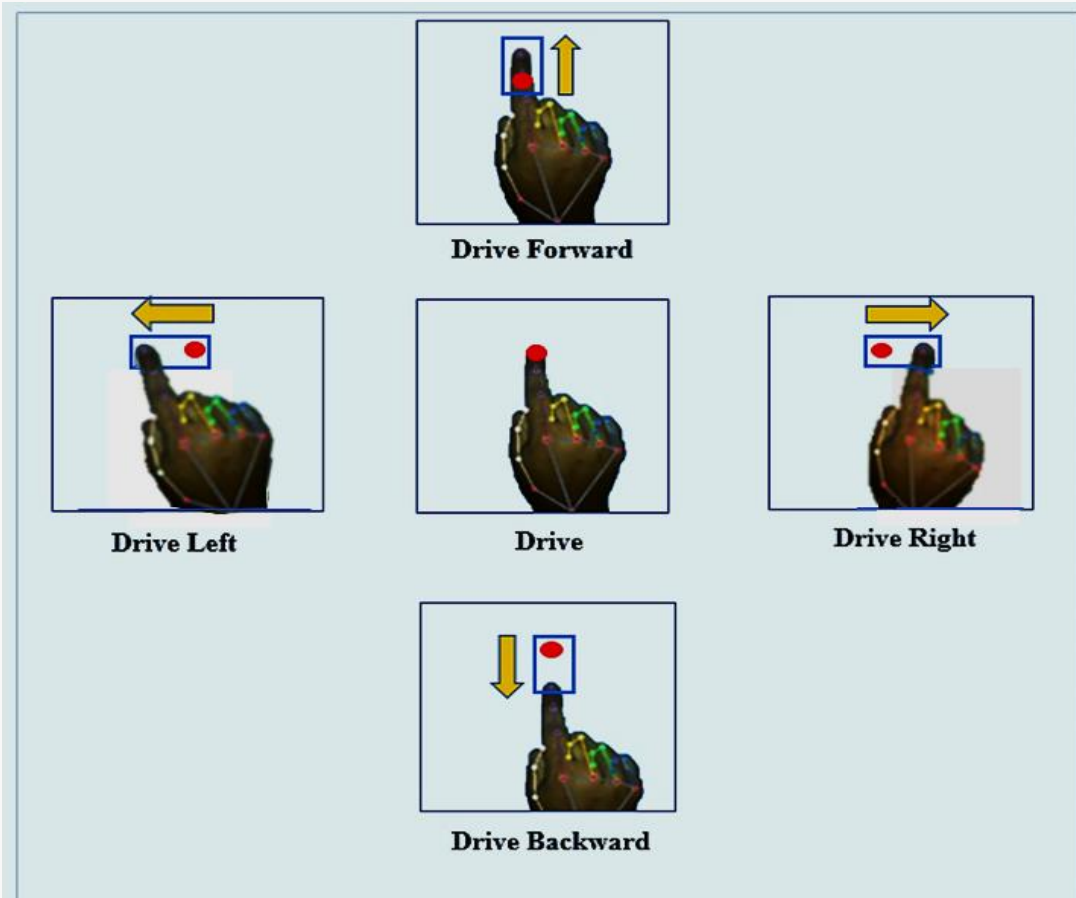


Figure 6. Gestures for basic and directional movements



Results

- ❑ The proposed system obtained a testing accuracy of 98.14% with a recognition rate of 21–27 FPS, with good Precision, Recall and F1 Score .
- ❑ This system is invariant of skin color, outdoor lighting environment and background.
- ❑ It's more user-friendly for disabled persons, as it works by short easy movement of fingers only.

Confusion matrix of the testing of the proposed hand gesture recognition system :

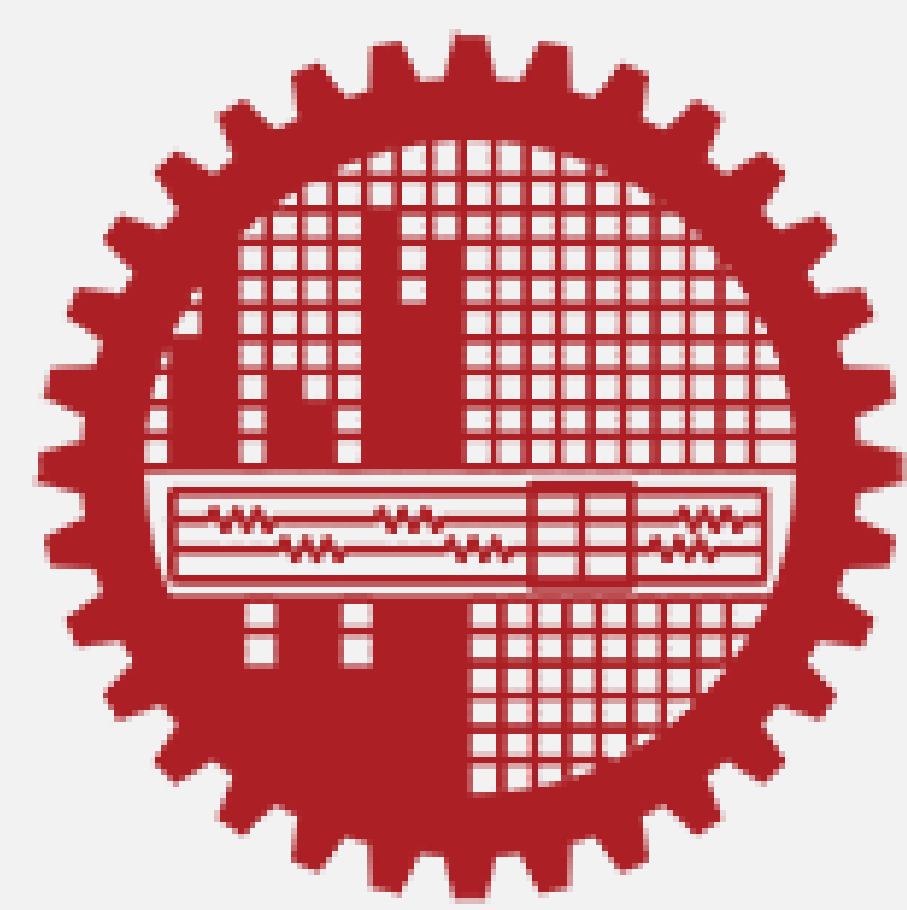
Recognized \ Performed	Stop	Horn	Drive Stop	Drive Forward	Drive Backward	Drive Left	Drive Right	Recall	Precision	F1 Score	Accuracy	Overall Accuracy
Stop	98							0.980	1.000	0.990	98%	98.14%
Horn		98						0.980	1.000	0.990	98%	
Drive Stop			99					0.990	0.900	0.947	99%	
Drive Forward				98				0.980	0.990	0.985	98%	
Drive Backward					99			0.990	1.000	0.995	99%	
Drive Left						97		0.970	0.990	0.980	97%	
Drive Right							98	0.980	0.990	0.985	98%	

Performance Comparison among different proposed systems:

Features	Proposed Systems			
	Gao et al.	Oliver et al.	Sadi, M.S. et al.	Proposed System
Functionality (Gesture Recognition)	Requires hand lifting, a Kinect Depth Camera and highly configured PC.	Requires wearing hand band, Joystick Manipulator, and accelerometer.	Requires movement of fingers, RGB Camera. It creates problems under daylight or if background matches with skin color.	Requires movement of fingers, RGB Camera. A mathematical model is used. Work well under daylight on any background.
Cost	High	Medium	Low	Low
Recognition Success rate	Depends on complexity (10–100%)	Actual accuracy is not measured	97.14%	98.14%

A Semi-Supervised Approach For Identifying Cyberbullying Text in Bangla Language

Aifa Faruque and Hossen A Mustafa



Abstract:

Social media has become increasingly popular, and it is now relatively simple to communicate with people online. As a result, there is more hate speech directed at specific individuals on social media. It is vital to create models that can help in automatically recognizing cyberbullying comments, which is becoming more and more prevalent on social media. Large-scale annotated corpora, which are currently uncommon in languages like Bangla, are typically needed for such models. However, it is exceedingly expensive and time-consuming to manually annotate corpora. To address this issue, we employ a semi-supervised self-training approach to replace the time-consuming human annotation process. We utilize Decision Tree, Logistic Regression, Random Forest, and Support Vector Machine for classification and experiment with combining different n-gram features to compare the effectiveness of these machine learning methods. In our experiment, we used the publicly available Bangla text dataset (44,001 comments) and experimental results show that the self-training algorithm with SVM is the base estimator that yields 90.57% F1-score.

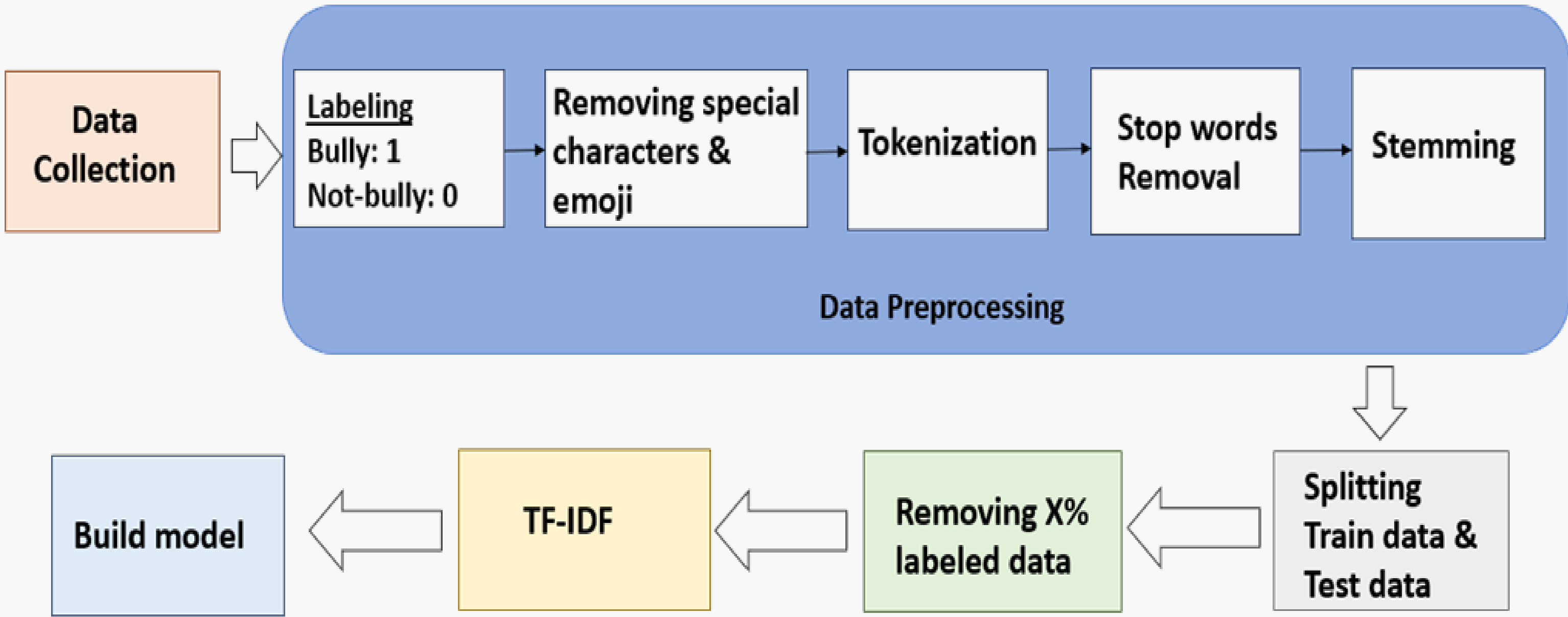
Background

- Saraiva et al. (2021) addressed the detection of toxic comments on a Portuguese language using a semi-supervised approach over a heterogeneous graph.
- In order to identify hate speech, Cahyana et al. (2022) suggested an automatic annotation procedure based on the idea of semi-supervised learning using the K-Nearest Neighbor algorithm utilizing feature extraction (TF-IDF).
- Shawon et. al. (2023) proposed semi-supervised GAN-LM architecture to classify Bengali fake reviews where they found that BanglaBERT with semi-supervised GAN obtained good accuracy.
- Ahmadinejad et al. (2023) propose a Machine Learning/Deep Learning-based approach for cyberbully detection and used balanced dataset for training the model and implementing a semi-supervised self-training algorithm for increasing the size of the labeled dataset.

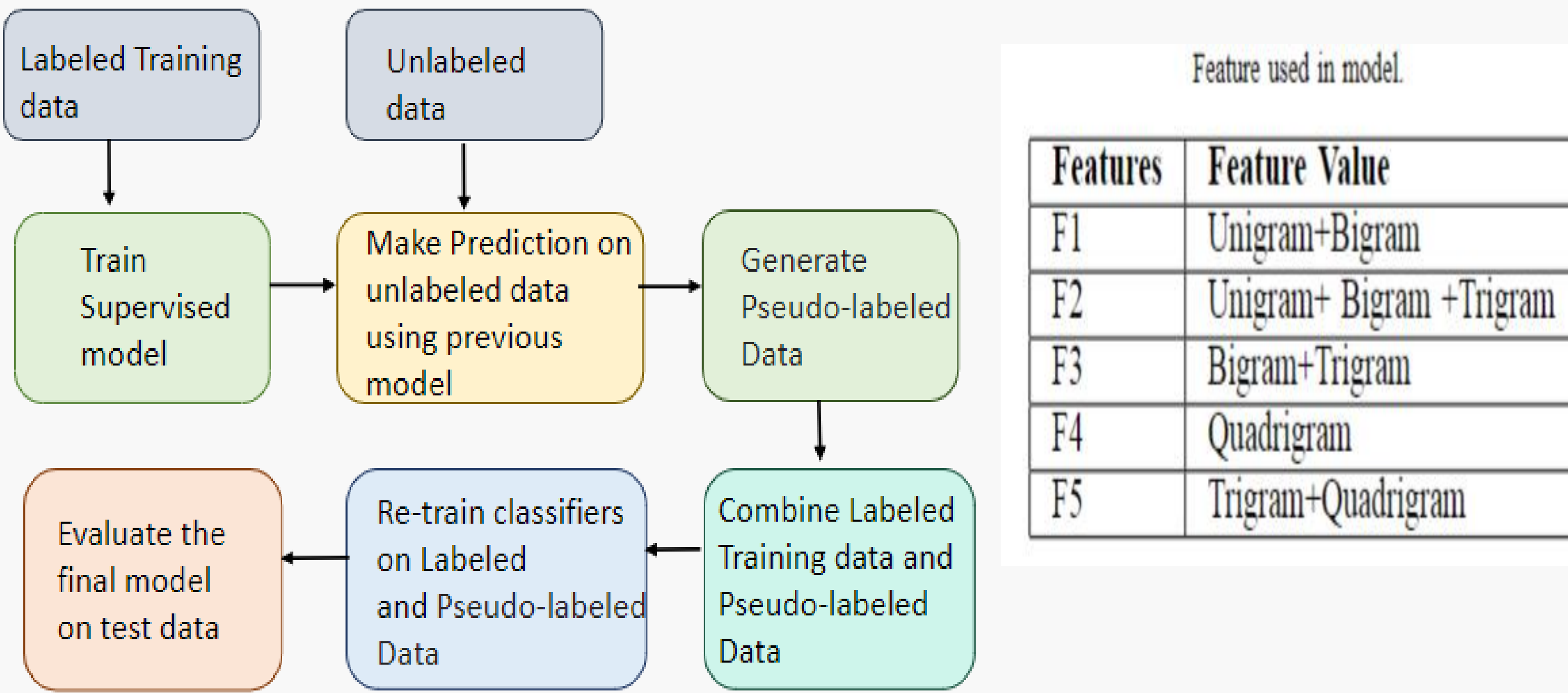
Motivation

- Self-training models use both labeled and unlabeled data, they belong to the Semi-Supervised branch of machine learning methods. Using the Self-Training model, a given supervised classifier (base estimator) can function as a semi-supervised classifier, allowing it to learn from unlabeled data. In each iteration, the base estimator predicts labels for the unlabeled samples and adds a subset of these labels to the labeled dataset.
- The majority of the research on cyberbullying identification on Bangla dataset has focused on supervised learning, which needs a large amount of textual data to be trained effectively.
- Semi-supervised approach may be effective in identifying cyberbullying texts on Bangla datasets.

Proposed Idea and Methodology



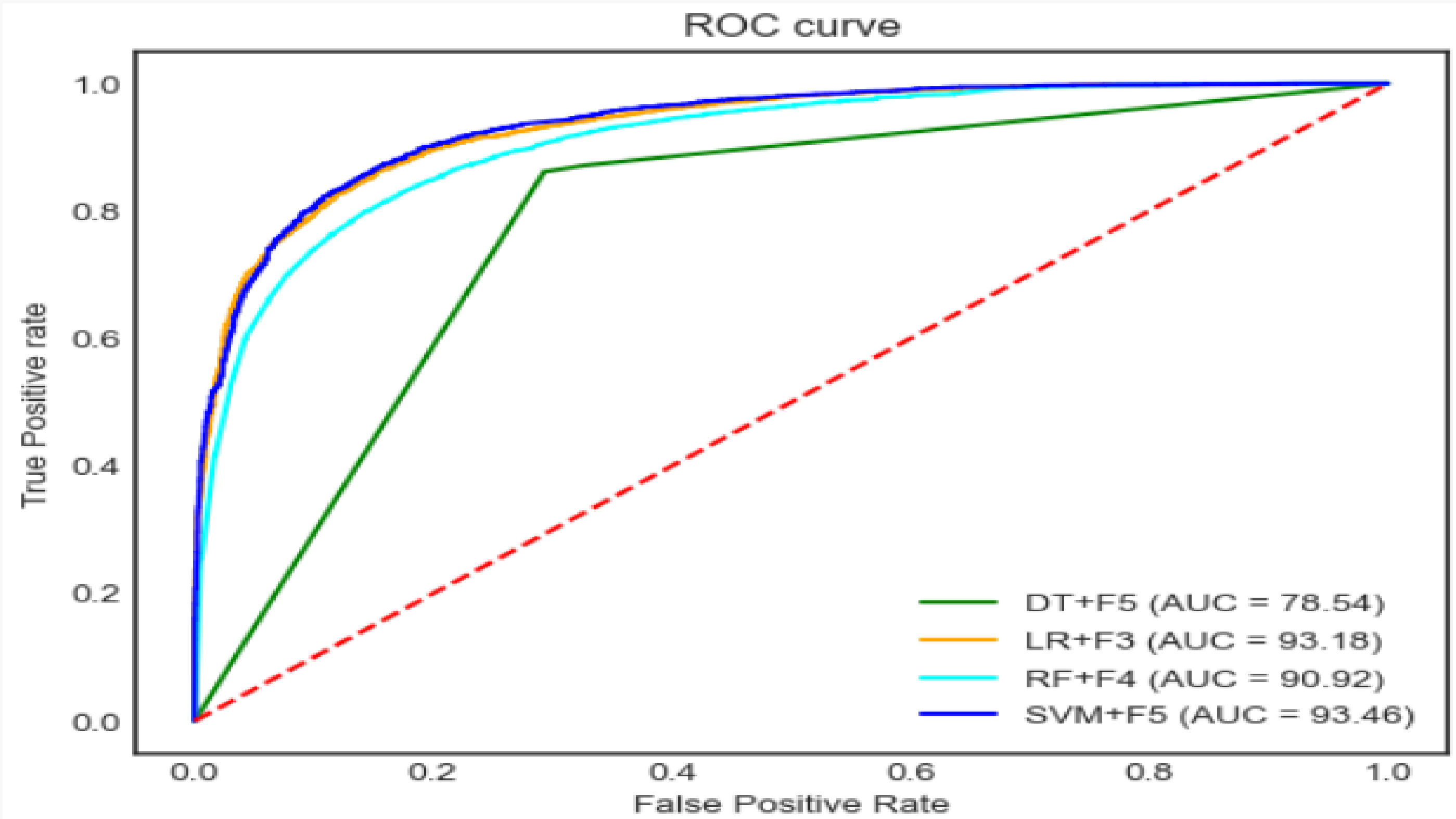
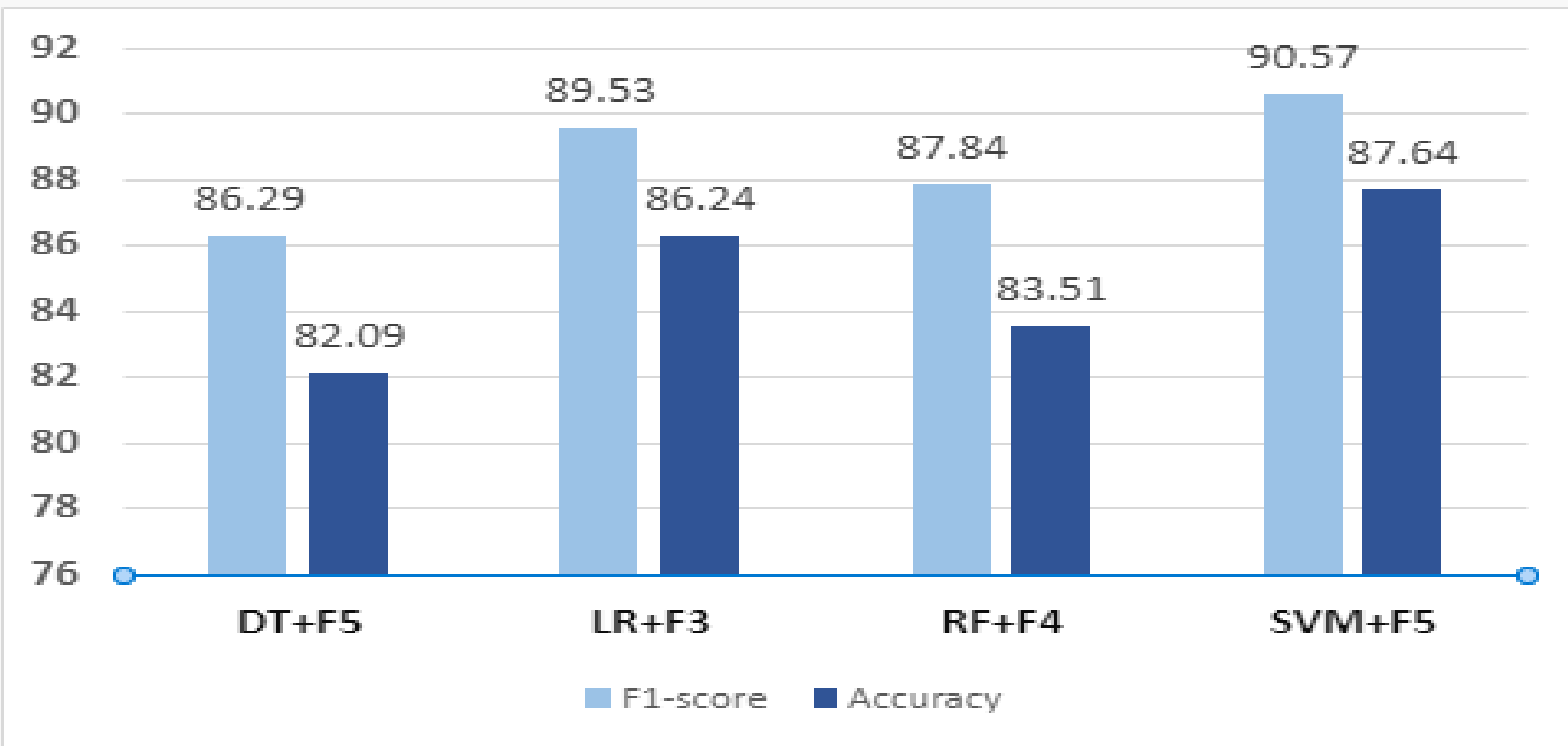
Self-training Method



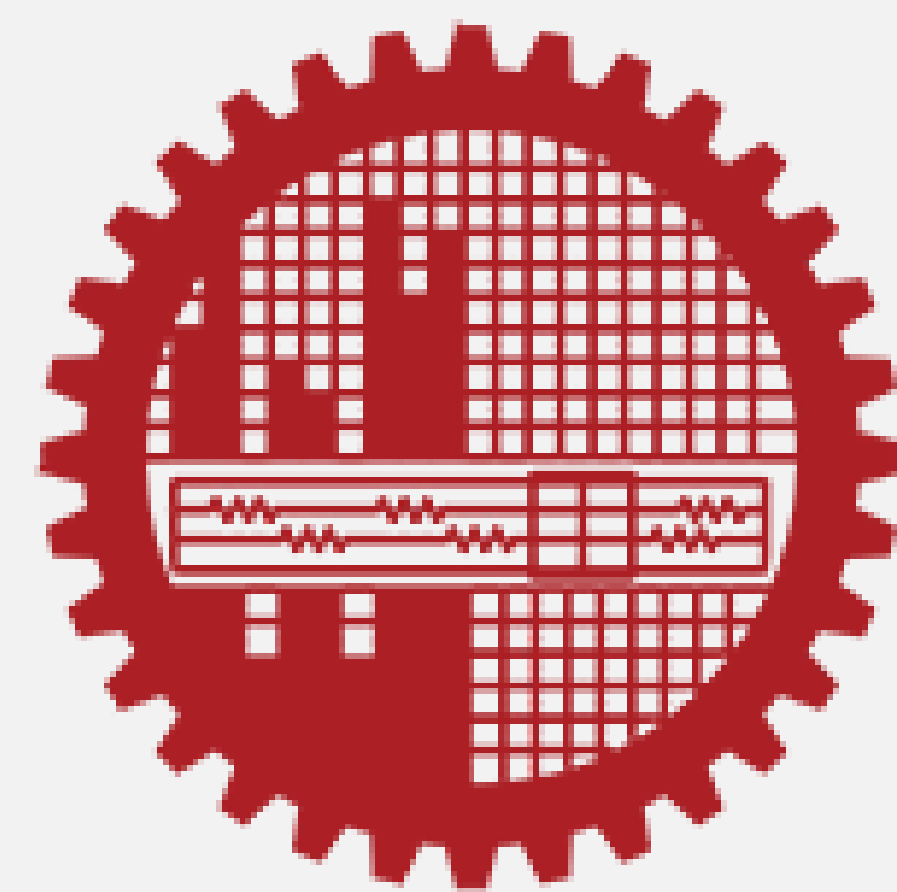
Feature used in model.	
Features	Feature Value
F1	Unigram+Bigram
F2	Unigram+ Bigram +Trigram
F3	Bigram+Trigram
F4	Quadrigram
F5	Trigram+Quadrigram

Results

Model		TF-IDF				
		F1	F2	F3	F4	F5
LR	A(%)	0.83	0.85	<u>0.86</u>	0.82	0.85
	FS(%)	0.87	0.89	0.89	0.87	0.88
RF	A(%)	0.75	0.78	0.78	<u>0.83</u>	0.82
	FS(%)	0.84	0.85	0.85	0.87	0.87
SVM	A(%)	0.76	0.85	0.86	0.86	<u>0.87</u>
	FS(%)	0.84	0.88	0.89	0.89	0.90
DT	A(%)	0.74	0.77	0.78	0.81	<u>0.82</u>
	FS(%)	0.8	0.82	0.83	0.85	0.86



DEVELOPMENT OF A LOW-COST EMBEDDED DATA ACQUISITION SYSTEM FOR CONCRETE RCP TEST.



Md. Mashiur Rahman, Dr. Md. Liakot Ali

Abstract

Data acquisition systems play a crucial role in various scientific and engineering applications, facilitating tasks such as data recording, storage, processing, and transmission. In the realm of Civil Engineering, one significant application is the Rapid Chloride Permeability Test (RCPT), which serves as an electrical indication of concrete's resistance to chloride ion penetration. This test is vital for durability-based quality control, aiding in the prediction of the service life of concrete structures. There are many automated data logging systems for RCPT testing on the market. Their high price makes it difficult for them to be widely used, particularly in small settings. For this reason, the demand for a reasonably priced substitute is high. This research aims to provide a reliable and accessible alternative to expensive automated data logging systems available in the market. Successful completion will yield a functional, cost-effective solution, addressing challenges such as noise cancellation, calibration accuracy, and hardware reliability. The outcomes will not only benefit concrete durability testing but also present opportunities for further research and development in the realm of automated data acquisition systems for various scientific and engineering applications.

Background & Motivation

During the RCPT, a constant voltage of around 60V or higher is applied to a concrete specimen for 6 hours, while the current passing through the concrete is recorded at 30-minute intervals to calculate coulombs. Manual data recording in this process is not only cumbersome for the responsible technician but also prone to human errors. Furthermore, it leads to a loss of productivity for the technician, who would otherwise need to be continuously engaged in the data recording process.

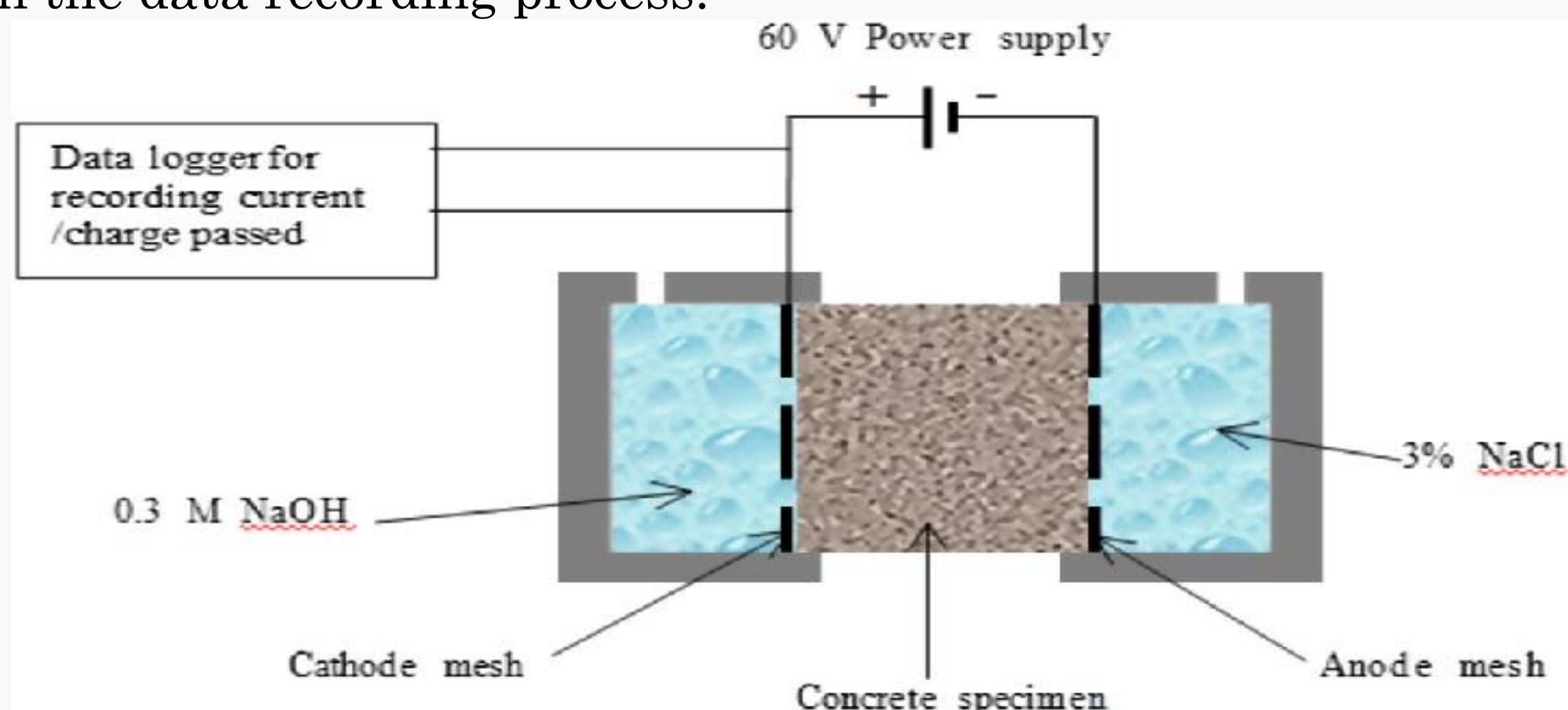


Figure-1: Systematic setup of RCPT test.

- Existing automated RCPT data logging systems are expensive, limiting their accessibility, particularly in local contexts. Addressing this issue, the proposed project aims to develop an affordable microcontroller-based data acquisition system tailored for concrete RCPT testing.
- This initiative aims to offer an affordable solution and presents opportunities for research and development. Overcoming challenges like sensor output noise and precise data logger calibration opens avenues for innovation. The project aims to contribute to advancing automated data acquisition systems, making them more accessible across scientific and engineering fields. Overall, the project aligns with the broader goal of improving efficiency, accuracy, and affordability in critical applications such as concrete durability testing.

Proposed Idea and Methodology

The simplified architecture of the proposed model for develop a microcontroller-based low-cost data acquisition system for concrete Rapid Chloride Permeability Testing (RCPT) is illustrated in Figure 2:

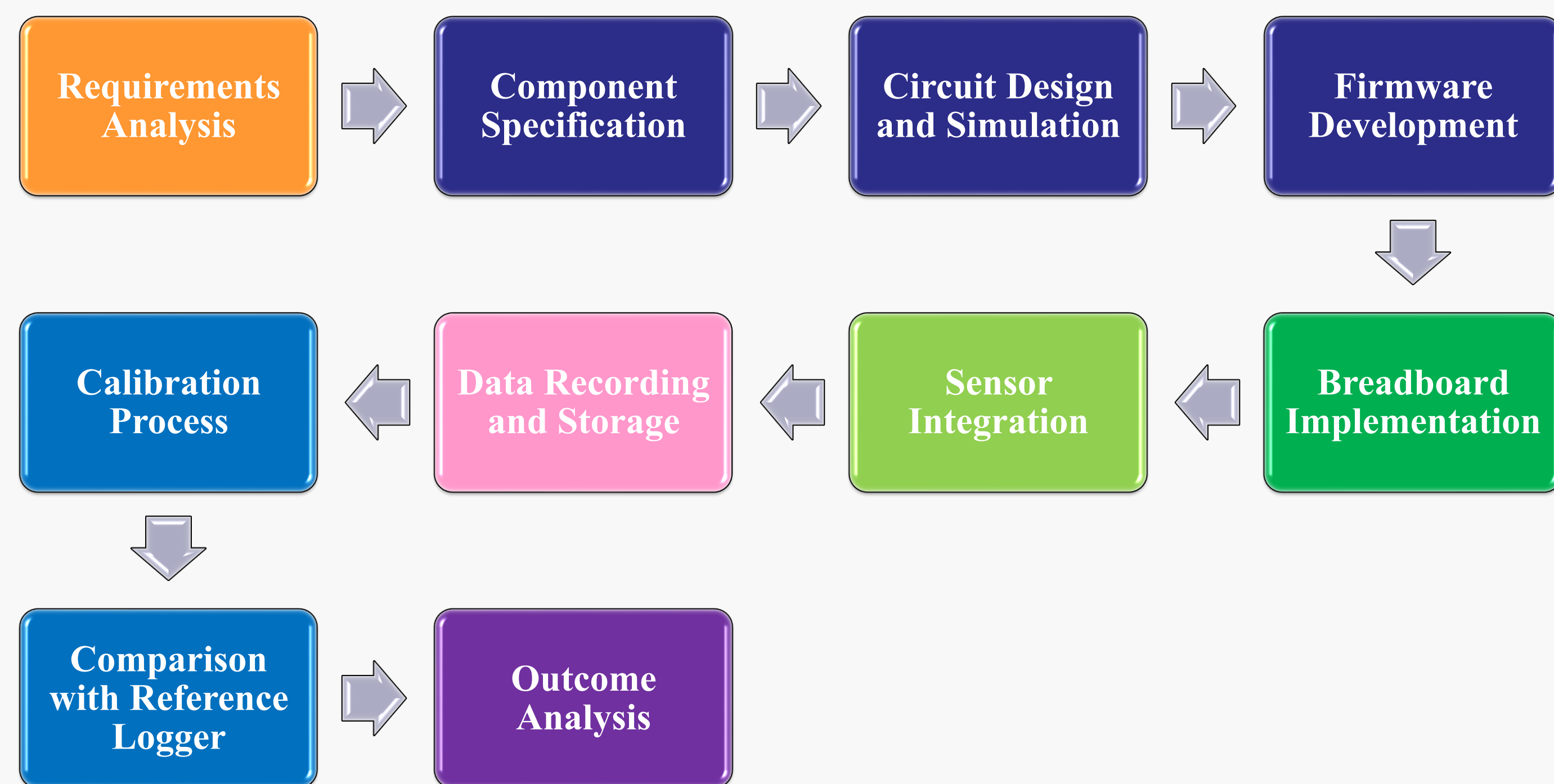


Figure-2: Simplified Conceptual Model of Proposed microcontroller-based low-cost data acquisition system for RCP test.

The methodology involves a systematic approach starting with a thorough analysis of project requirements. Following this, specific hardware and software components, including the 'Voltage sensor,' 'Current sensor,' and 'SD Card module,' are carefully selected for optimal accuracy. The circuit is designed and simulated using Proteus software, and firmware is developed using Code Vision software for an AVR microcontroller. The circuit is then implemented on a breadboard, integrating sensors for voltage and current measurements during the 6-hour Rapid Chloride Permeability Test (RCPT). Data is recorded and stored on an SD card at 30-minute intervals. Calibration processes are incorporated for enhanced accuracy, comparing measured values with those obtained from a digital multimeter. This comprehensive methodology aims to deliver a cost-effective and reliable microcontroller-based data acquisition system tailored for concrete RCPT testing.

Results

Expected Results:

- Functional, cost-effective data acquisition system for concrete durability testing.
- Accurate and consistent measurement of voltage and current during the RCPT.
- Successful comparison with a standard reference data logger, validating reliability.

Significance:

- Cost-effective alternative for researchers and technicians in concrete durability testing.
- Potential for broader applications in various scientific and engineering scenarios.

Challenges:

- Noise cancellation from sensor outputs.
- Calibration accuracy for voltage and current.

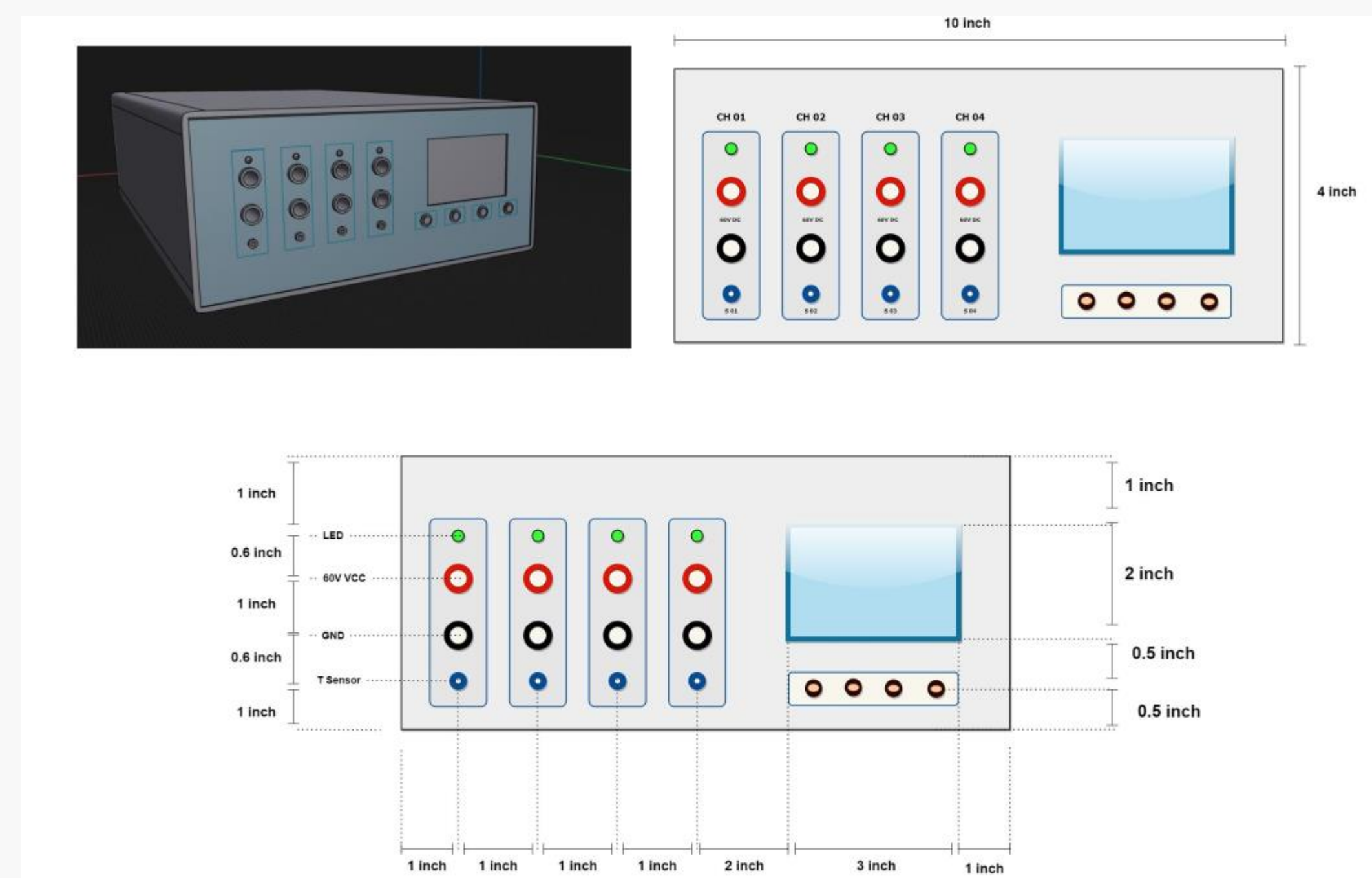
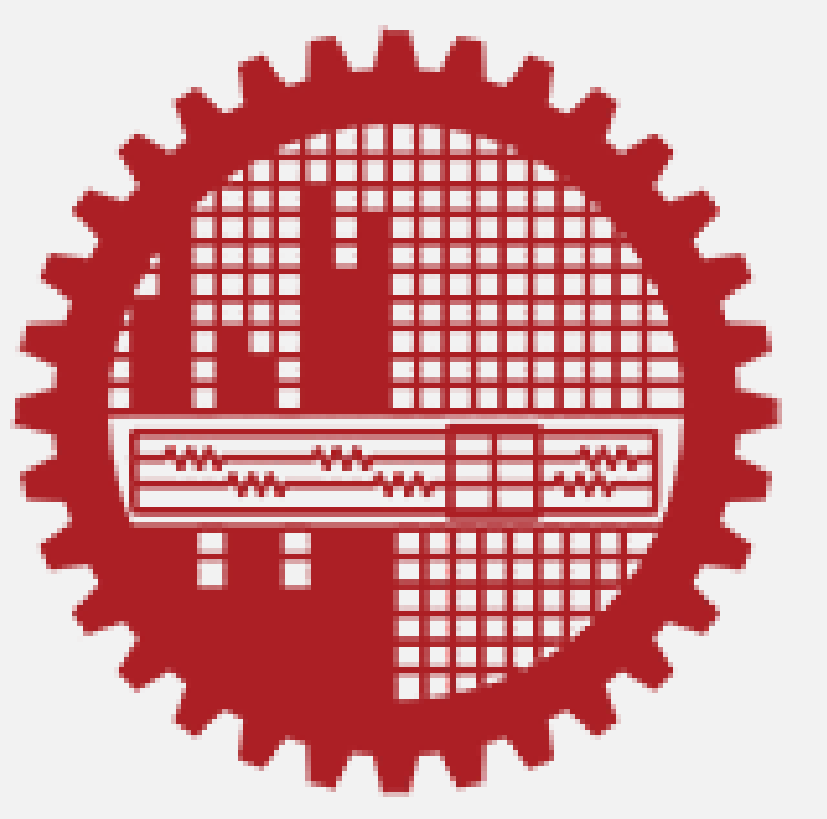


Figure-3: Front View of RCPT Body Model.

PREDICTING PRETERM BIRTH DURING COVID-19 ERA: A STUDY IN BANGLADESH



Nadia Binte Asif and Dr. Hossen A Mustafa

COVID-19 & BANGLADESH

In Bangladesh, the **COVID-19** pandemic started in **March 2020**, and still, cases are being found of this deadly virus. In this developing country with **poor health** conditions and **sanitation**, it is essential to understand how this deadly virus impacted the lives of people, especially **pregnant women**.

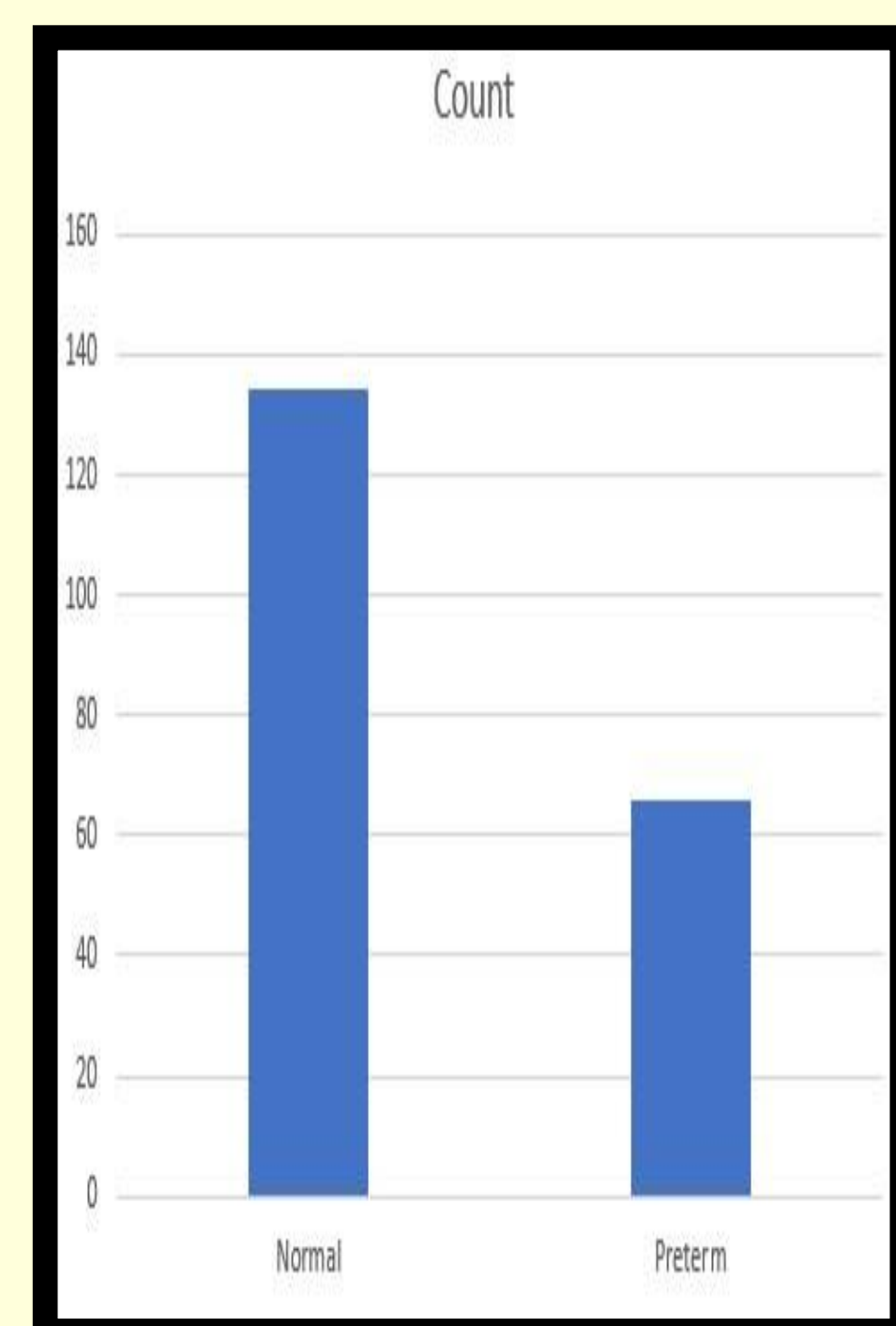
PURPOSE

Machine learning approaches have previously been proposed to **predict preterm birth**; however, the **association of COVID-19** with other comorbidities is a fairly new term, and very few studies have been conducted on the premise of Bangladesh.

BACKGROUND

- Harvey et al. (2021) from **Tennessee** established a **correlation between COVID-19** and other co-morbidities with preterm birth, and got **95% accuracy**
- Perez et al. (2021) from **Spain** also proved using **multivariate logistic regression** that affected pregnant women are more likely to give birth to **preterm baby**.
- In Bangladesh, the study of Masud et al. (2021) showed **95% confidence interval**;
- Algorithms:** Firth Logistic Regression, Binary Linear Regression and Multivariate Regression

IMBALANCED CLASS



SOLUTION

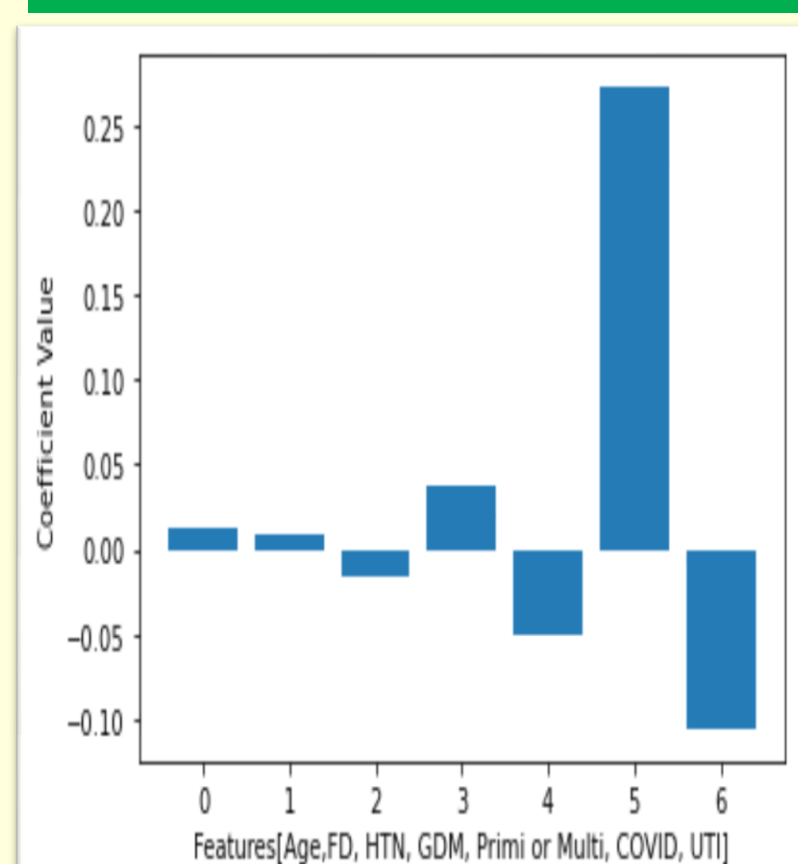
SMOTE-N

FEATURE IMPORTANCE

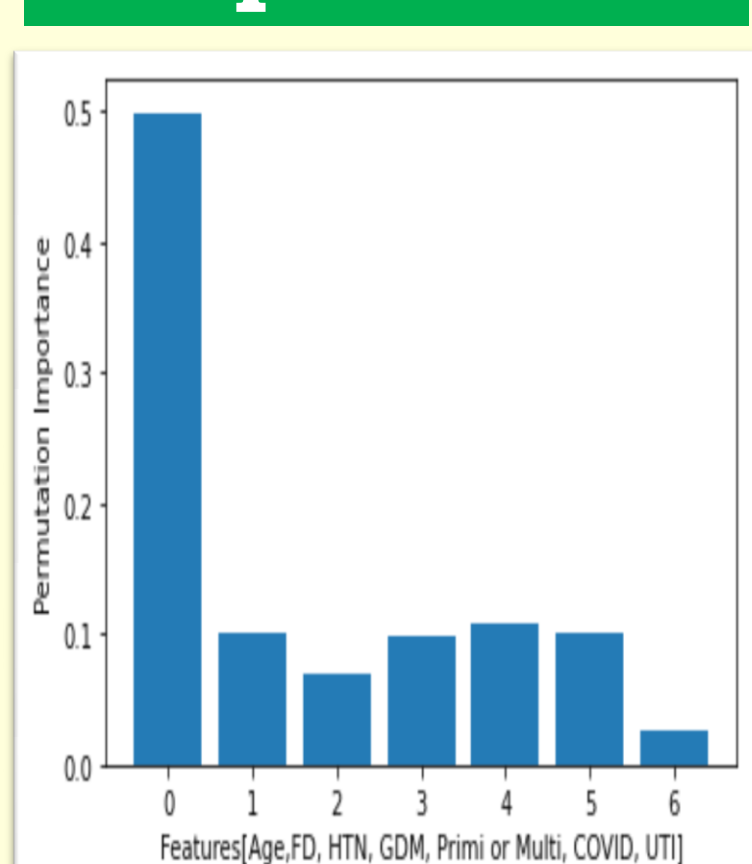
A large number of independent variables or features may lead to overfitting of the model and increase the overall cost. Obtained dataset is an **imbalanced** one and appropriate features are selected based on the following table:

Features	Linear Regression	Logistic Regression	Random Forest	XGB Regressor	SVM
AGE	0.01227	0.06165	0.50440	0.12859	2
FD	0.00914	0.09796	0.09517	0.12658	0
HTN	-0.01556	-0.00688	0.06662	0.17535	6
GDM	0.03813	0.14663	0.09278	0.13304	4
Primi or Multi	-0.05015	-0.21164	0.10846	0.12927	1
COVID	0.27246	1.00855	0.10513	0.18724	3
UTI	-0.10601	-0.30223	0.02739	0.11992	5

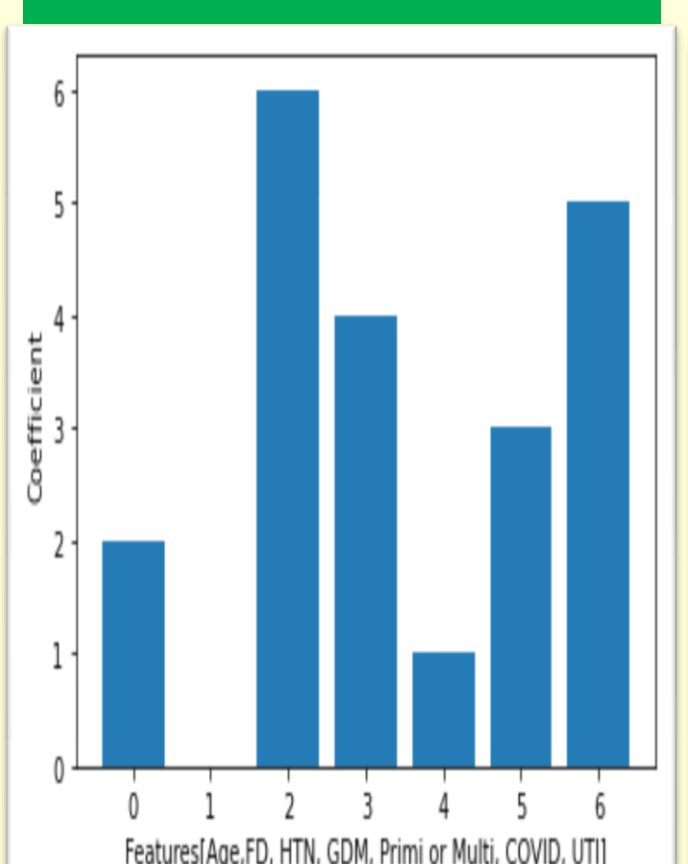
Regression Coefficients



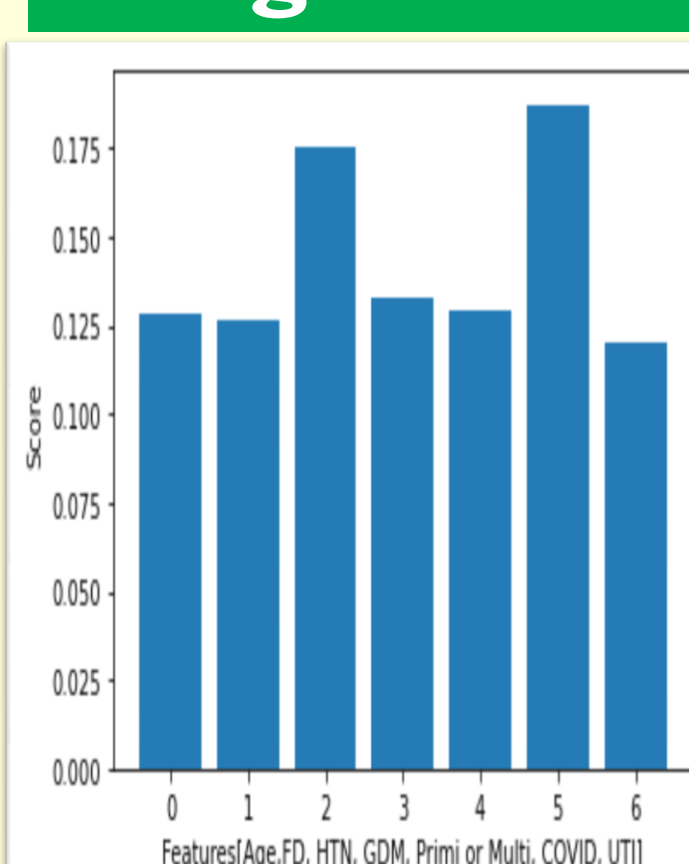
Permutation Importance



SVM



XGB Regressor

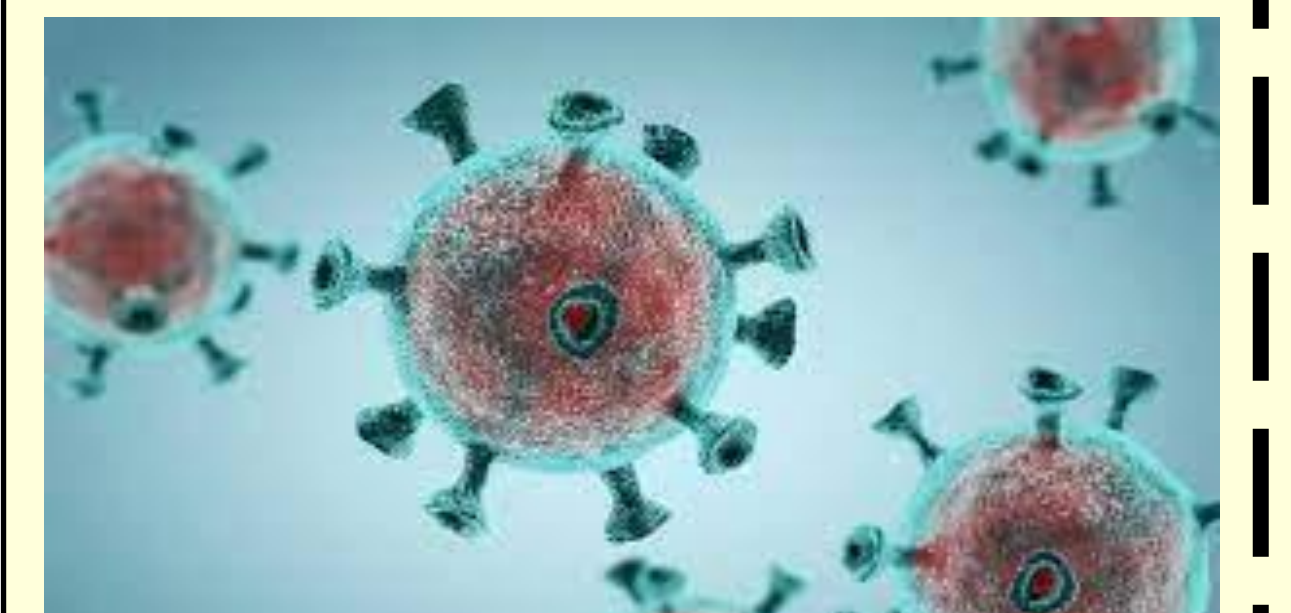


WHAT IS PRETERM BIRTH ?

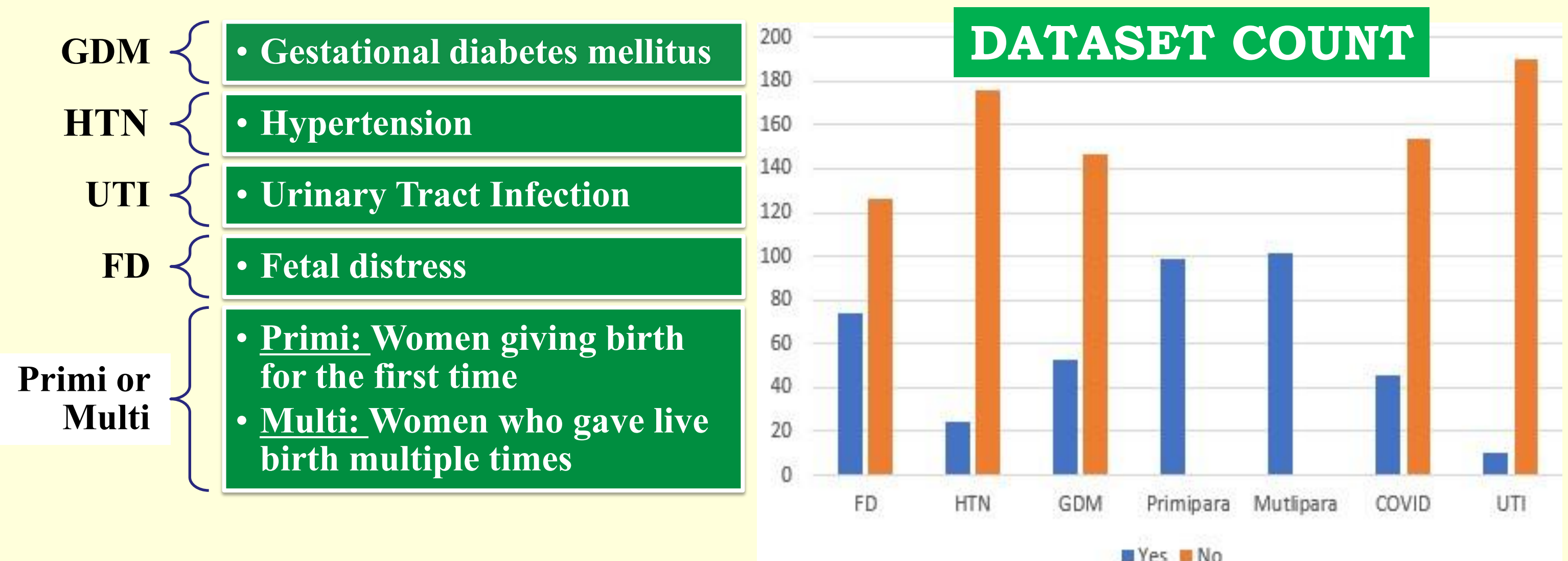
- A fetus that is born prior to 37 weeks.**
- The normal period for a healthy fetus to be born is 39 to 40 weeks.
- This periodic time is divided into three phases: First (12 weeks), Second (13 to 27 weeks), Third trimester (28 to 40 weeks).

HYPOTHESIS

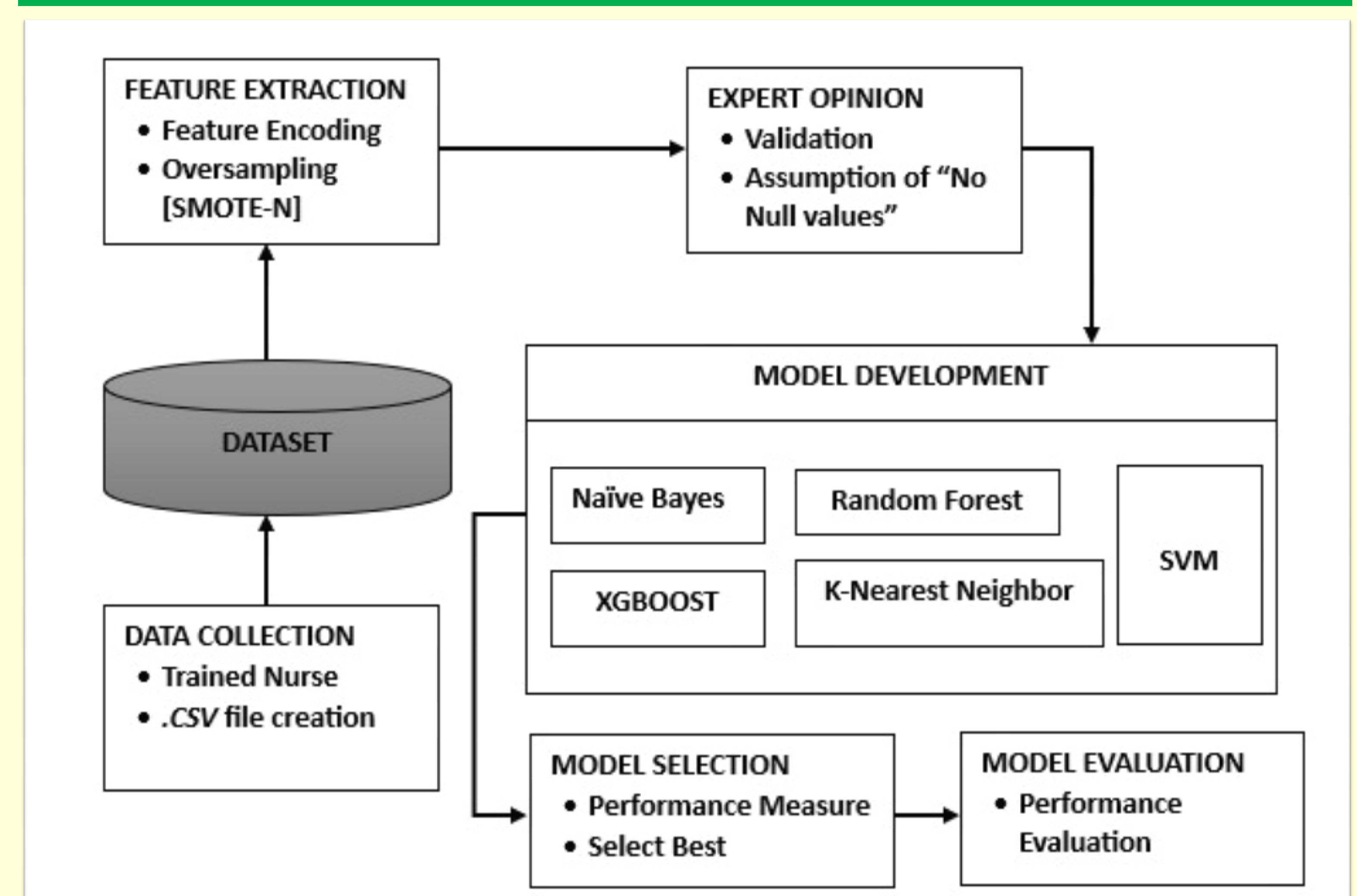
If a pregnant woman suffers from COVID-19 and also suffers from any single co-morbidities like Fetal Distress (FD), Urinary tract infection (UTI), Hypertension (HTN), or Gestational Diabetes Mellitus (GDM), chances are more that she will give birth to a preterm baby.



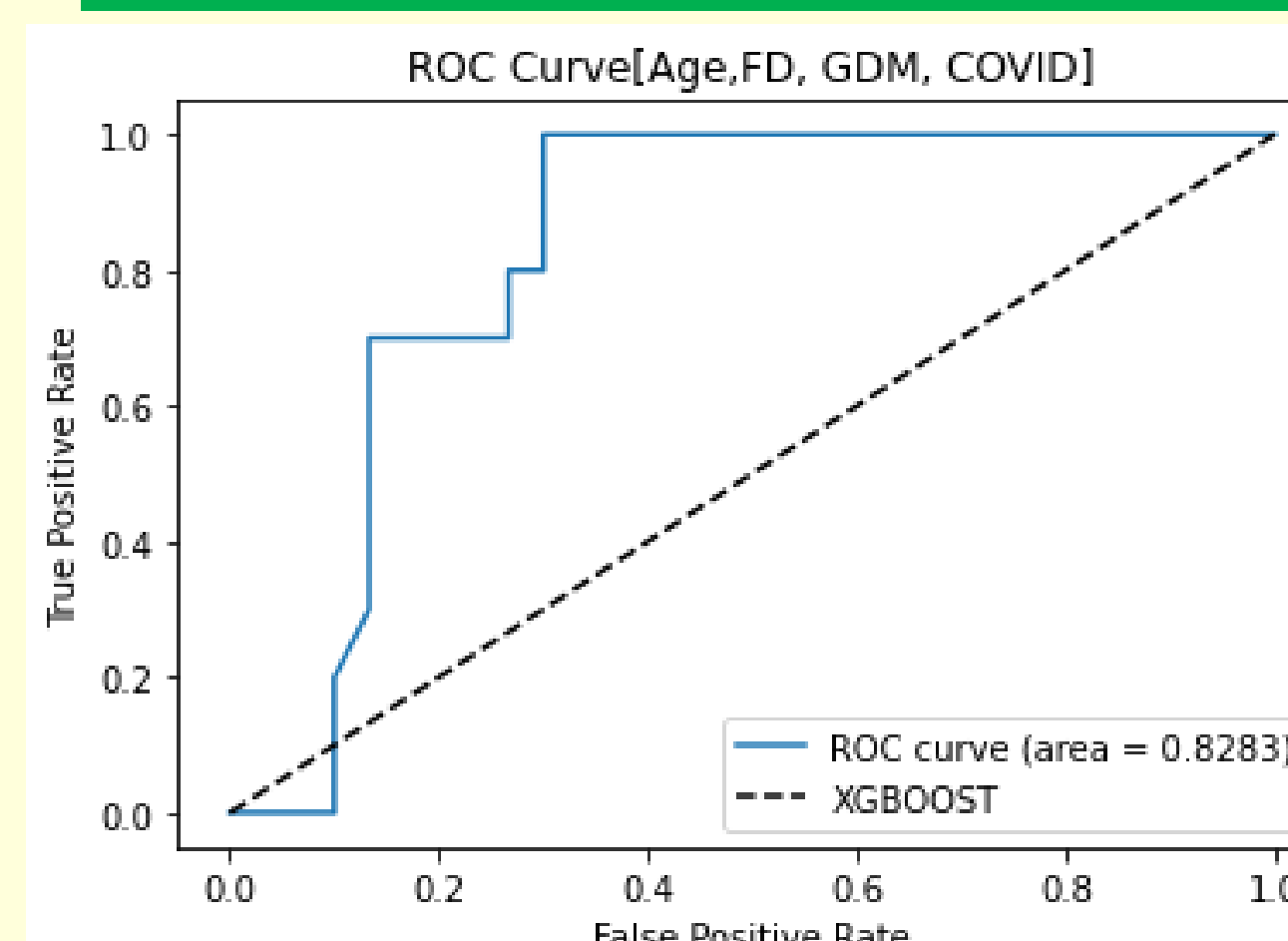
DATASET COUNT



MODEL WORKFLOW



AUR_ROC CURVE



CLASSIFIER SUMMARY

Classifier	Accuracy	AUC_ROC Score
Naïve Bayes	0.7500	0.7184
SVM	0.7500	0.7569
RF	0.7500	0.5865
XGBOOST	0.7500	0.8283
KNN	0.7000	0.7115

UPSHOT

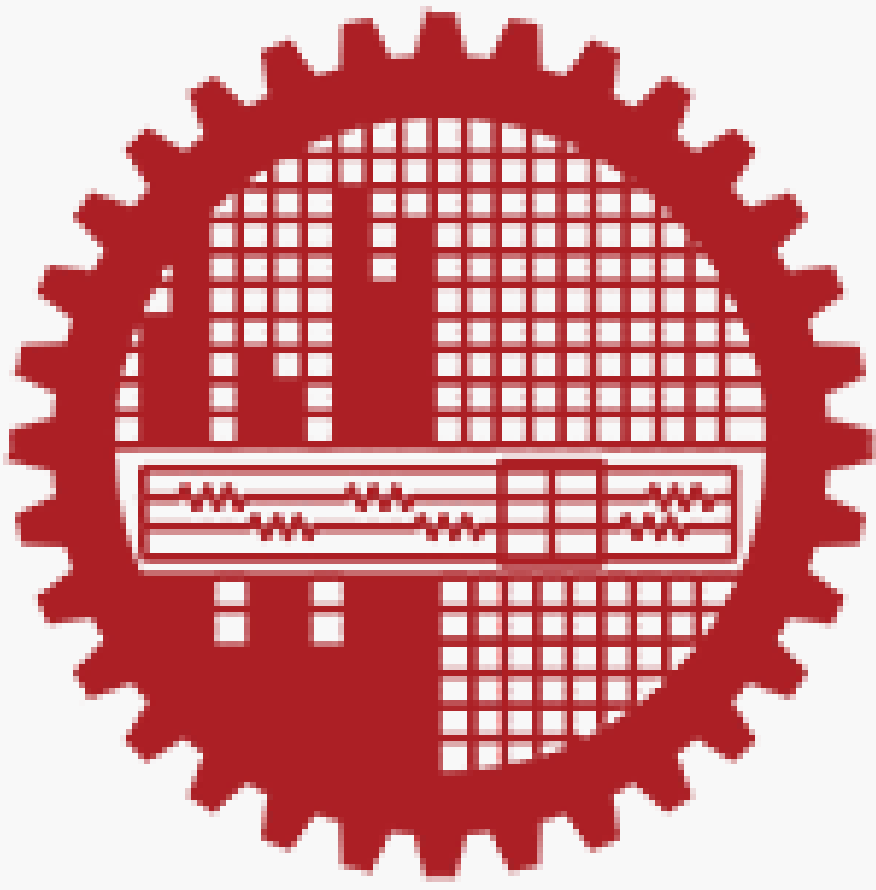
- Machine learning** can be used to establish a **direct relationship** between COVID-19 and Preterm Birth
- The **predictive model** will help the **doctors** to make **informed decision** when an affected woman comes to their care.

RESULT

- Best Classifier: XGBOOST**
- Accuracy: 75%**
- ROC Score: 82.83%**
- Precision: 0.700**
- F1-Score: 0.8077**
- Sensitivity: 0.9545**
- Specificity: 0.5000**

Title: Cyber Threat Detection in Computer Networks: Leveraging Machine Learning for Enhanced Security

Md. Zobair Raihan and Dr. Md. Saiful Islam



Abstract

As the reliance on computer networks continues to grow, so does the complexity of cyber threats that pose a substantial risk to information security. This thesis explores the application of machine learning techniques for cyber threat detection in computer networks, aiming to enhance overall security. Leveraging the power of advanced algorithms, the research focuses on the development of a robust intrusion detection system capable of identifying diverse cyber threats.

The study employs machine learning models to analyze network traffic patterns, identify anomalies, and promptly detect potential security breaches. A comprehensive investigation is conducted, evaluating the effectiveness of various machine learning algorithms in enhancing the security posture of computer networks. Specifically, the analysis incorporates the use of Analysis of Variance (ANOVA) to assess the significance of different features in the context of cyber threat detection. This statistical technique aids in identifying crucial variables contributing to the detection accuracy, thereby optimizing the model's performance.

Furthermore, the research integrates the Mutual Information algorithm, a powerful information-theoretic approach, to measure the dependency between variables and uncover hidden relationships within the network data. By leveraging Mutual Information, the thesis aims to enhance the accuracy and reliability of the machine learning models, ensuring a more robust and adaptive cyber threat detection system.

The outcomes of this research contribute to the advancement of cyber threat detection methodologies, emphasizing the importance of machine learning techniques and statistical analyses in fortifying the security infrastructure of computer networks. The findings not only showcase the efficacy of the proposed models but also provide insights into the optimal utilization of ANOVA and Mutual Information algorithms for enhanced security in the dynamic landscape of cyber threats.

Background & Motivation

In recent years, the ubiquity and complexity of computer networks have exponentially increased, making them indispensable for communication, commerce, and various critical operations. However, this widespread connectivity has also led to a surge in cyber threats, ranging from sophisticated malware to targeted attacks, jeopardizing the integrity and confidentiality of sensitive information. Consequently, the imperative to fortify computer networks against evolving cyber threats has become a paramount concern for both organizations and individuals.

Traditional security measures, such as firewalls and signature-based intrusion detection systems, are struggling to keep pace with the agility and sophistication of modern cyber threats. As a result, there is a pressing need for innovative approaches that can adapt to the dynamic nature of cyber threats and proactively identify malicious activities before they compromise the integrity of computer networks.

The motivation for this thesis stems from the recognition that machine learning presents a transformative solution to the challenges posed by contemporary cyber threats. Leveraging the power of machine learning algorithms offers the potential to enhance the security posture of computer networks by enabling intelligent, adaptive, and real-time threat detection. By harnessing the ability of these algorithms to learn from patterns in vast datasets, it becomes possible to identify anomalies, recognize emerging threats, and respond swiftly to mitigate potential risks.

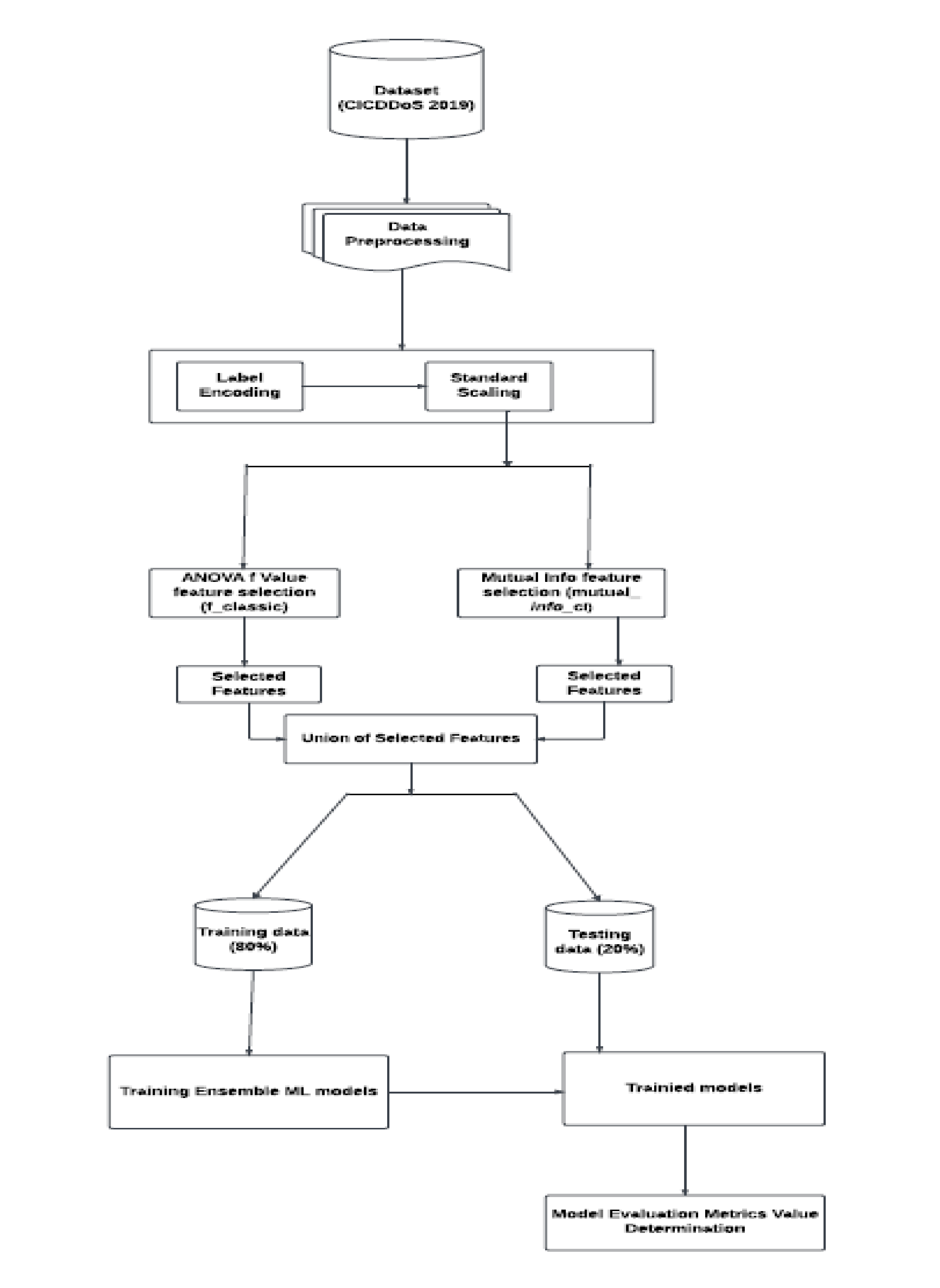
Furthermore, the motivation behind incorporating machine learning in cyber threat detection is rooted in the quest for a more robust and proactive security framework. Unlike traditional rule-based systems, machine learning models can evolve and adapt over time, learning from new data and continuously improving their ability to discern between normal and malicious network behaviour.

This research aims to contribute to the advancement of cyber threat detection methodologies by systematically exploring and evaluating the application of machine learning techniques

Proposed Idea and Methodology

The proposed idea is to enhance DDoS attack detection by combining multiple machine-learning techniques. This involves investigating the current state-of-the-art in DDoS attack detection methods, evaluating different machine learning algorithms, and developing a framework that combines multiple learning methods to improve the effectiveness of detecting DDoS traffic.

The methodology of this research is given below:

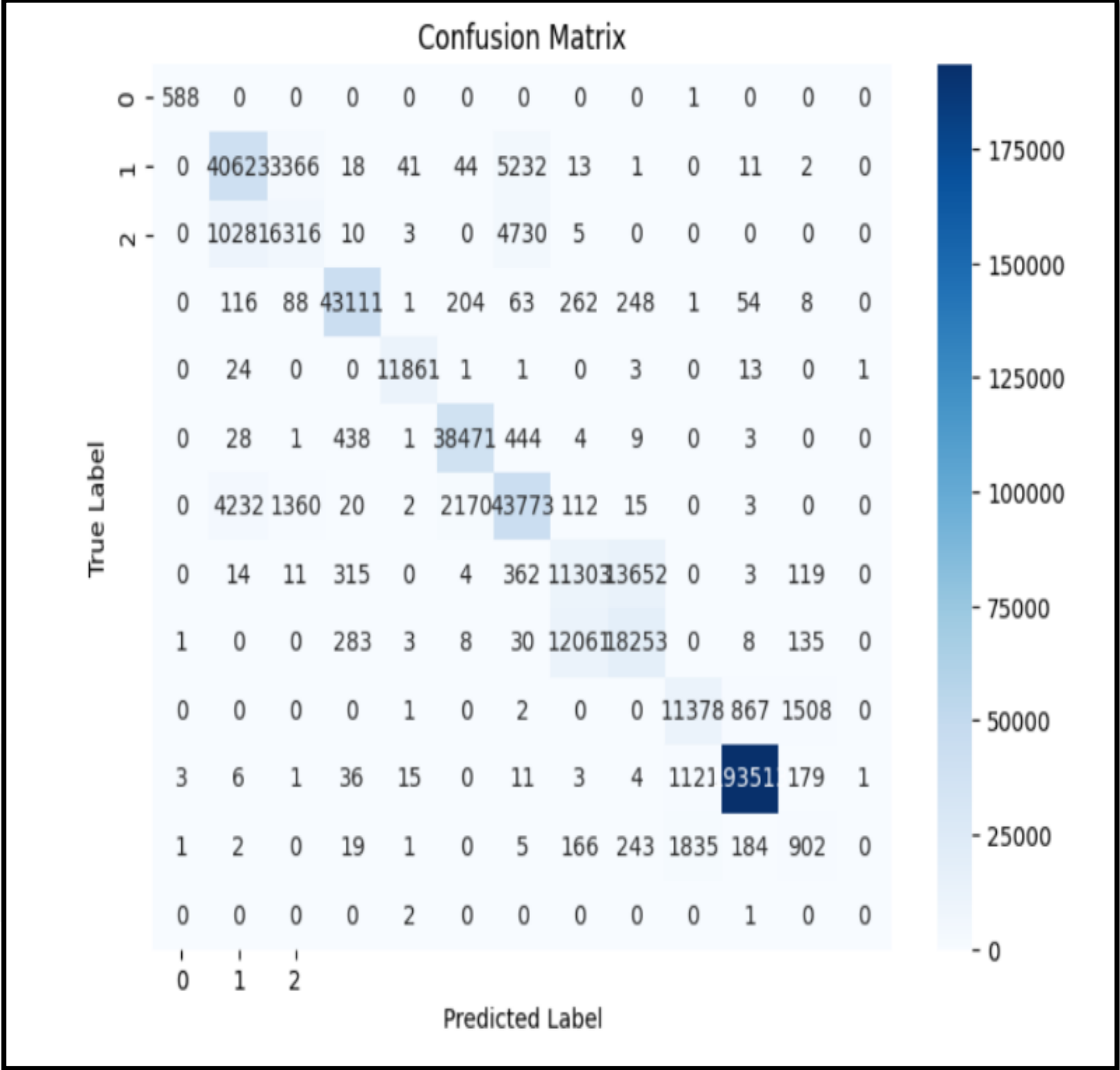
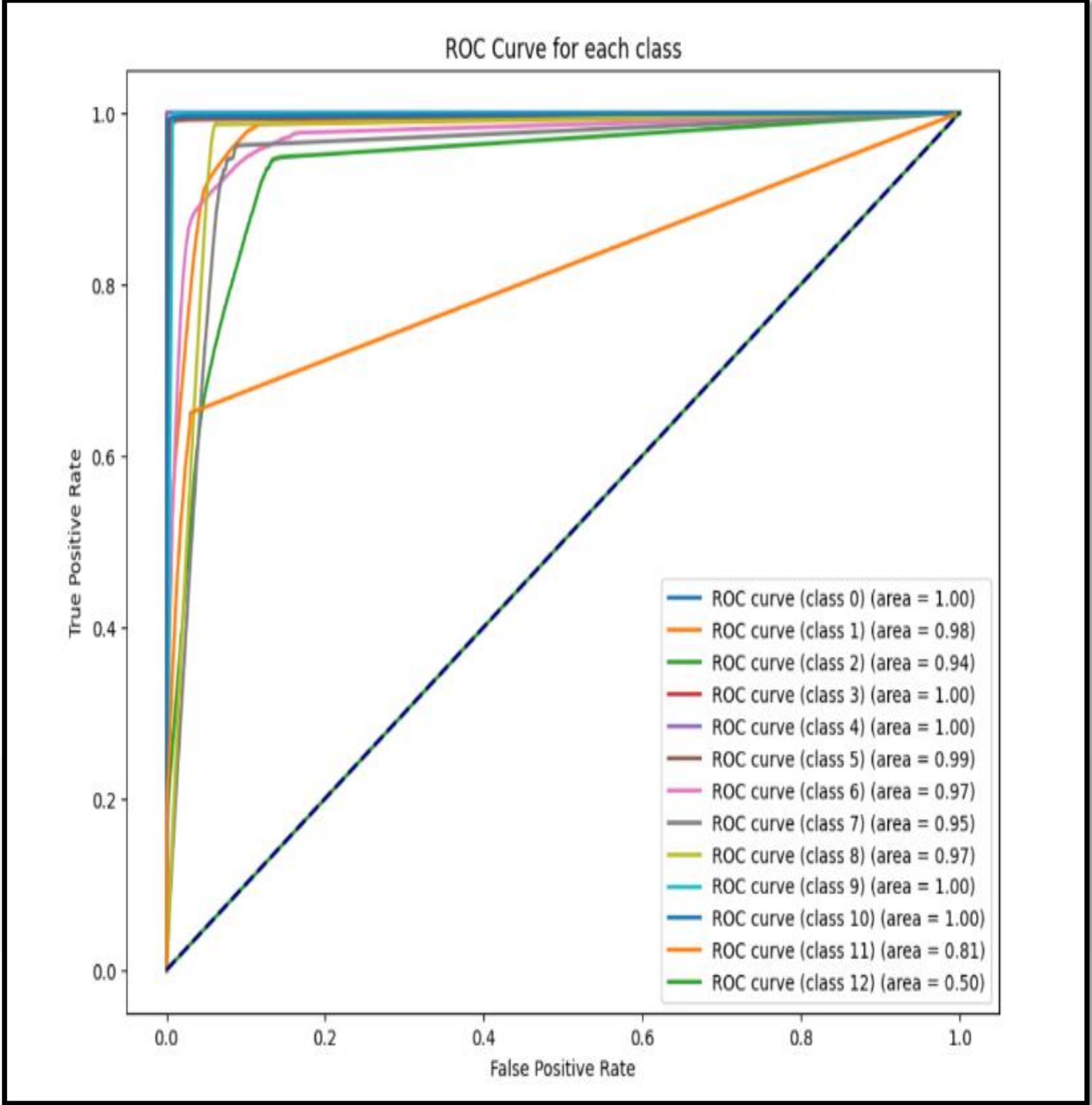


Results

- The proposed model, which uses 30 features, can accurately and successfully detect the DDoS traffic in (0-10) classes. The result of the various evaluation metrics is given below in figures.
- The accuracy of the proposed RF model is 86%, the error rate is 0.14

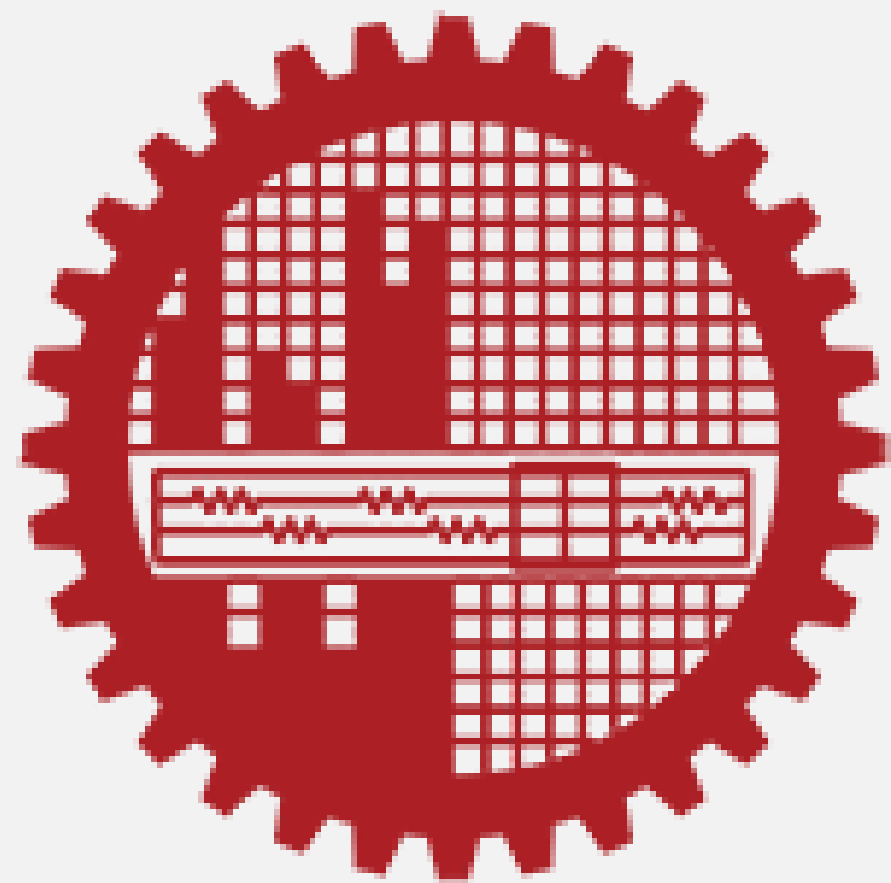
Accuracy: 0.8626012821197275

Classification Report:				
	precision	recall	f1-score	support
0	0.99	1.00	0.99	589
1	0.73	0.82	0.78	49351
2	0.57	0.30	0.39	21345
3	0.97	0.98	0.98	44156
4	0.99	1.00	1.00	11904
5	0.94	0.98	0.96	39399
6	0.80	0.85	0.82	51687
7	0.47	0.44	0.45	25783
8	0.56	0.59	0.58	30782
9	0.79	0.83	0.81	13756
10	0.99	0.99	0.99	194893
11	0.32	0.27	0.29	3358
12	0.00	0.00	0.00	3
accuracy			0.86	487006
macro avg	0.70	0.69	0.70	487006
weighted avg	0.86	0.86	0.86	487006



CREATION OF NOVEL MASHUP DATA AND DETECTION OF CATEGORICAL WEB ATTACKS

Shakil Ahammad and Hossen A Mustafa



Abstract

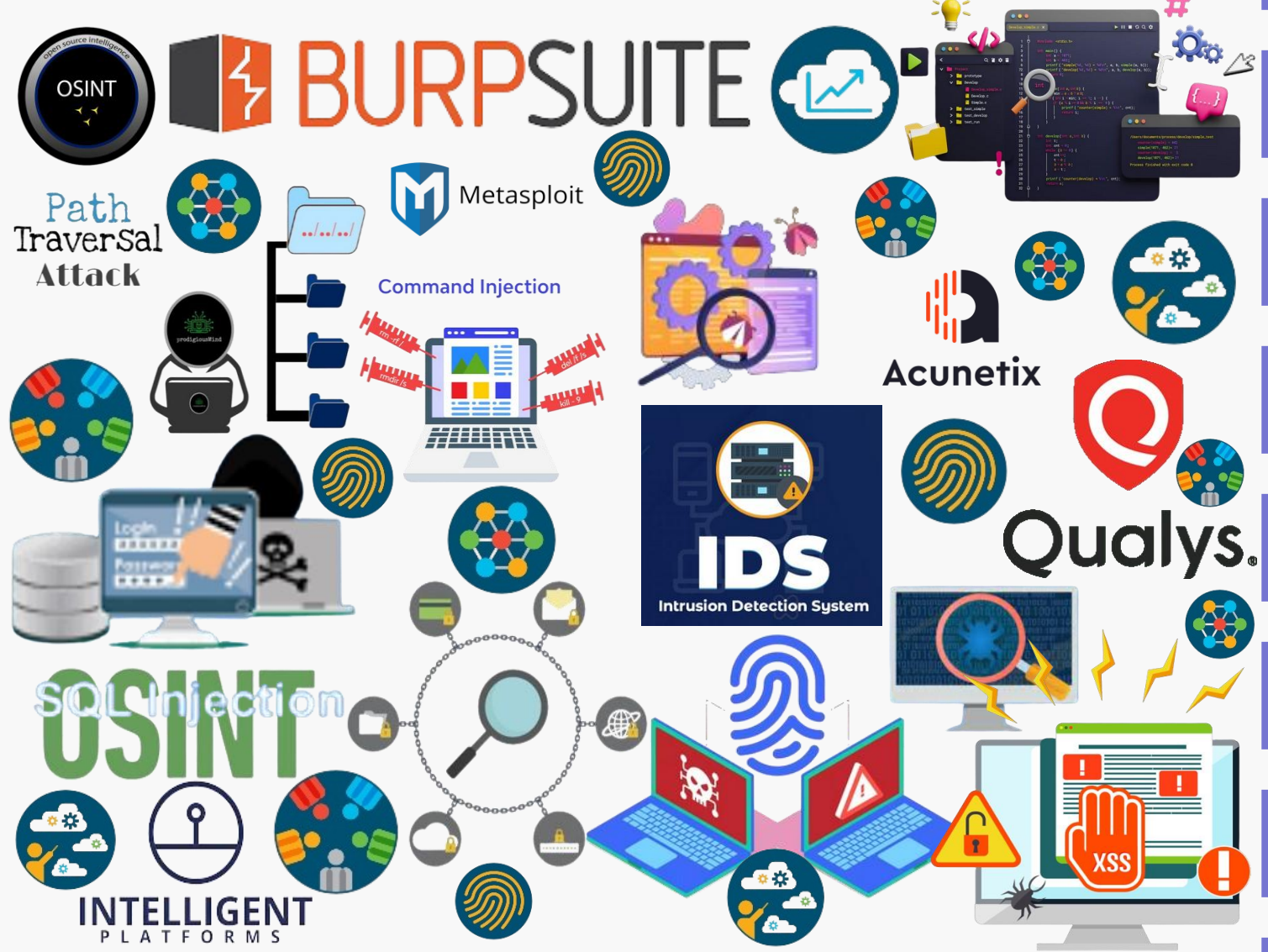
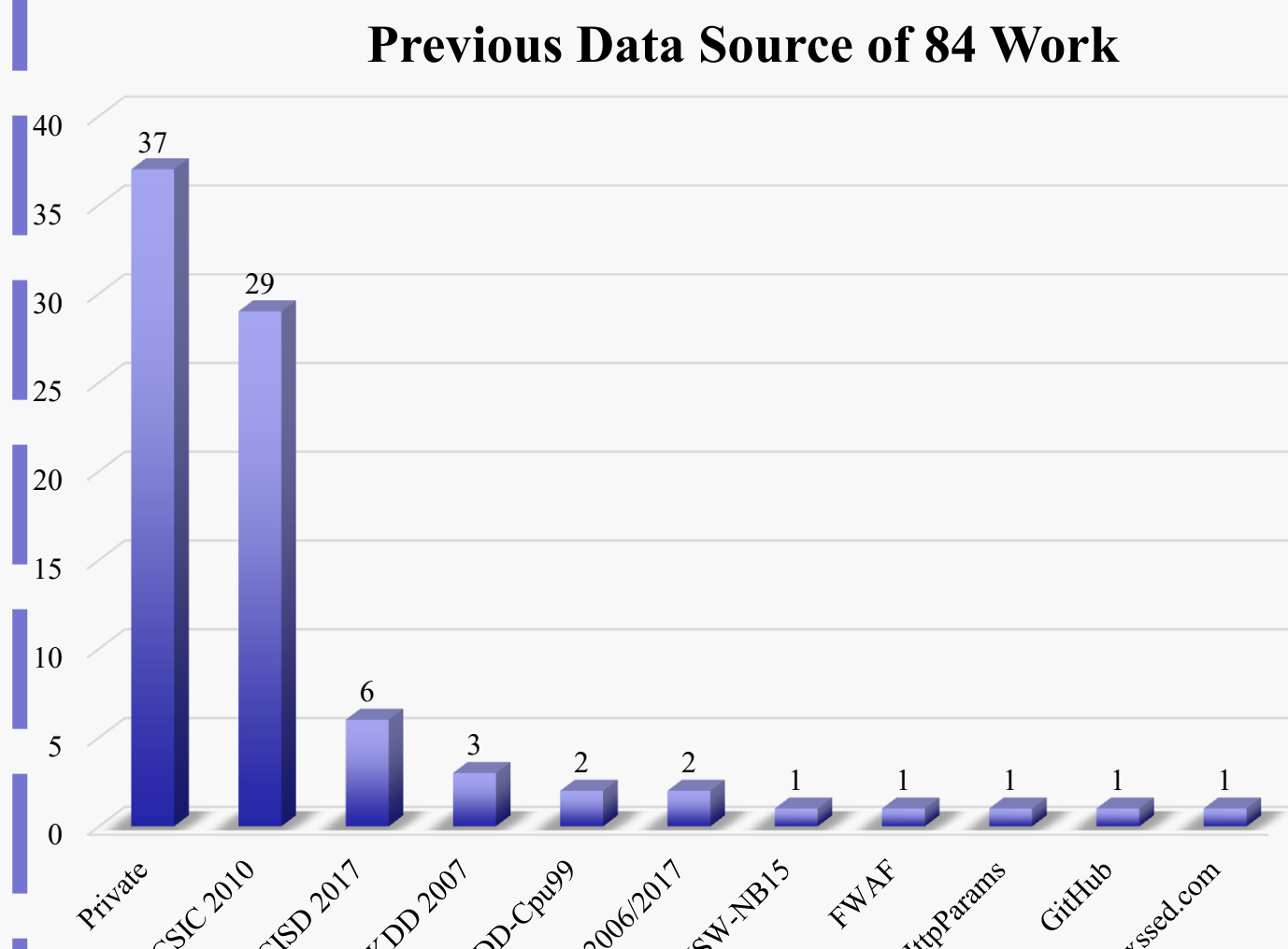
Information is considered the next-generation asset and by utilizing this component, people can be more dynamic as smart residents of the global village in this universe. A sustainable cyberspace is treated as the key element to ensure the accessibility of information. To design a sustainable cyberspace, a host-based intrusion detection system (IDS) is considered a main component for tracking any intruder's activity. So, considering the dimension of this information contamination approach [attack vector], security researchers are always facing challenges to ensure a resilient IDS to ensure security. Analyzing the attacking surface, sustainable cyberspace design and cost-effective IDS have been expected to deploy massively. In this context, there would be a scope to enhance the capability of a host-based IDS engine considering a Novel Mashup Data concept where data from different heterogeneous sources are used to build the IDS based on machine learning models. Early experimental results show that our proposed system can effectively detect different types of attacks using mashup data.

Background

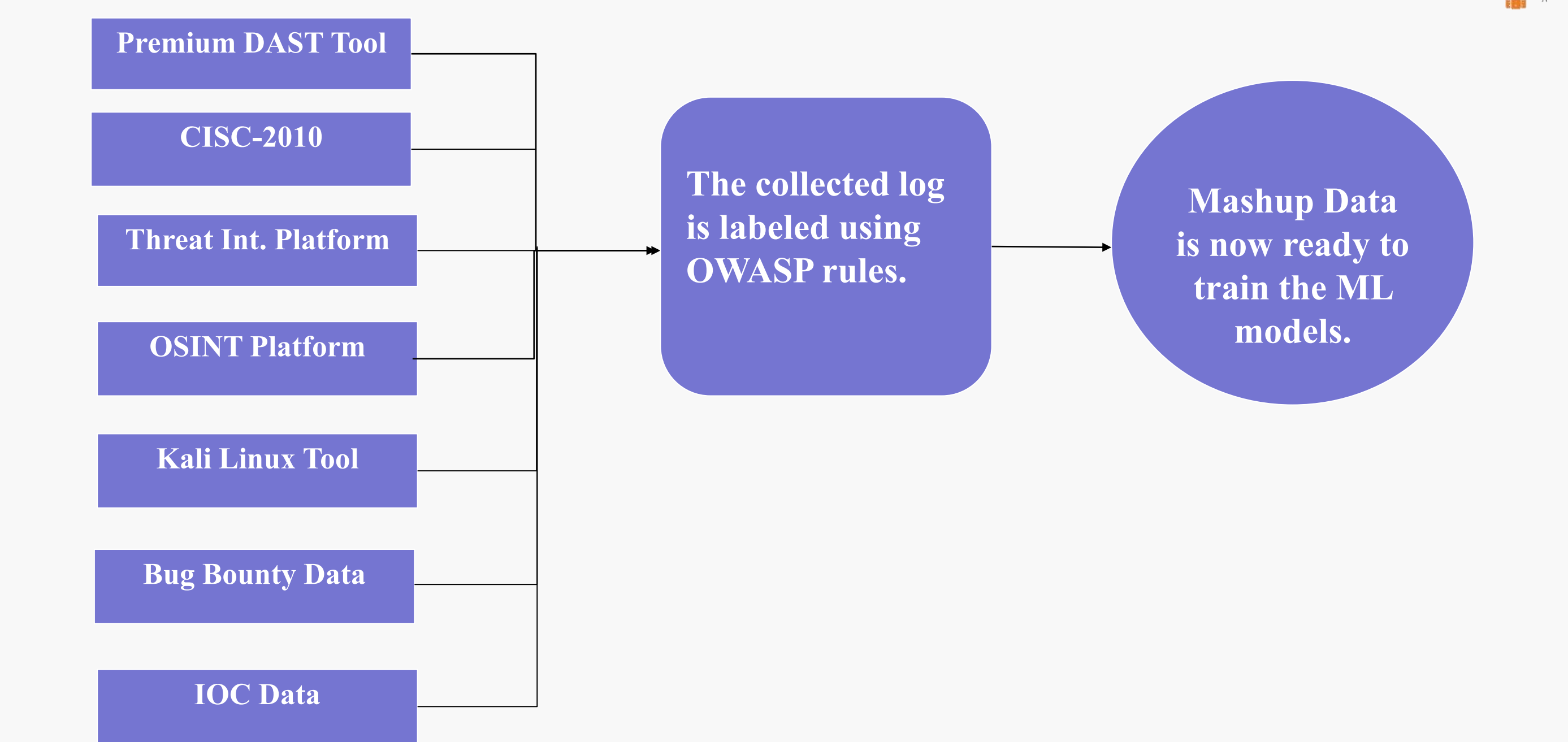
- Rokia Lamrani Alaoui (2022) presented more than 84 works as a literature review, where the Web Vulnerabilities Countermeasures are described as a picture of the approach to designing IDS engine.
- Branislav Raji'c (2022) proposed, using the Web application scanning process (DAST), a model for early web attack detection the intruder's activity.
- S.S. Anandha Krishnan (2021) presented SQLi detection using ML, This work has been explained using the GitHub data source to increase performance of the ML models..
- Yuqi Yu (2020) proposed a model of deep-HTTP traffic analysis process, where the Anomaly HTTP Traffic Detection using ML models comparison is described.

Motivation

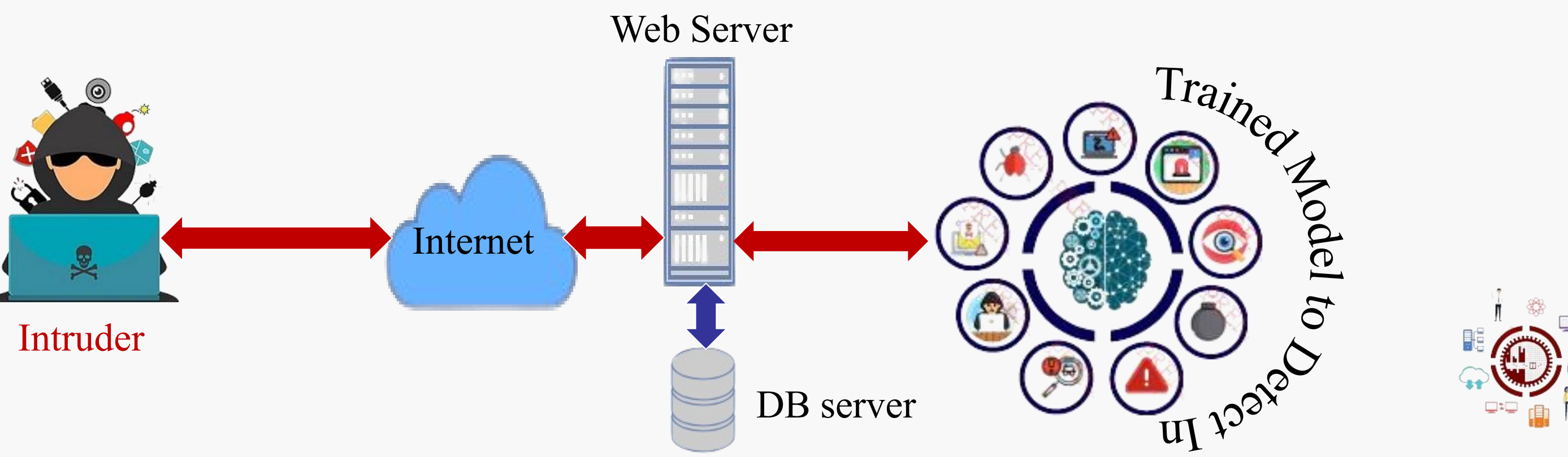
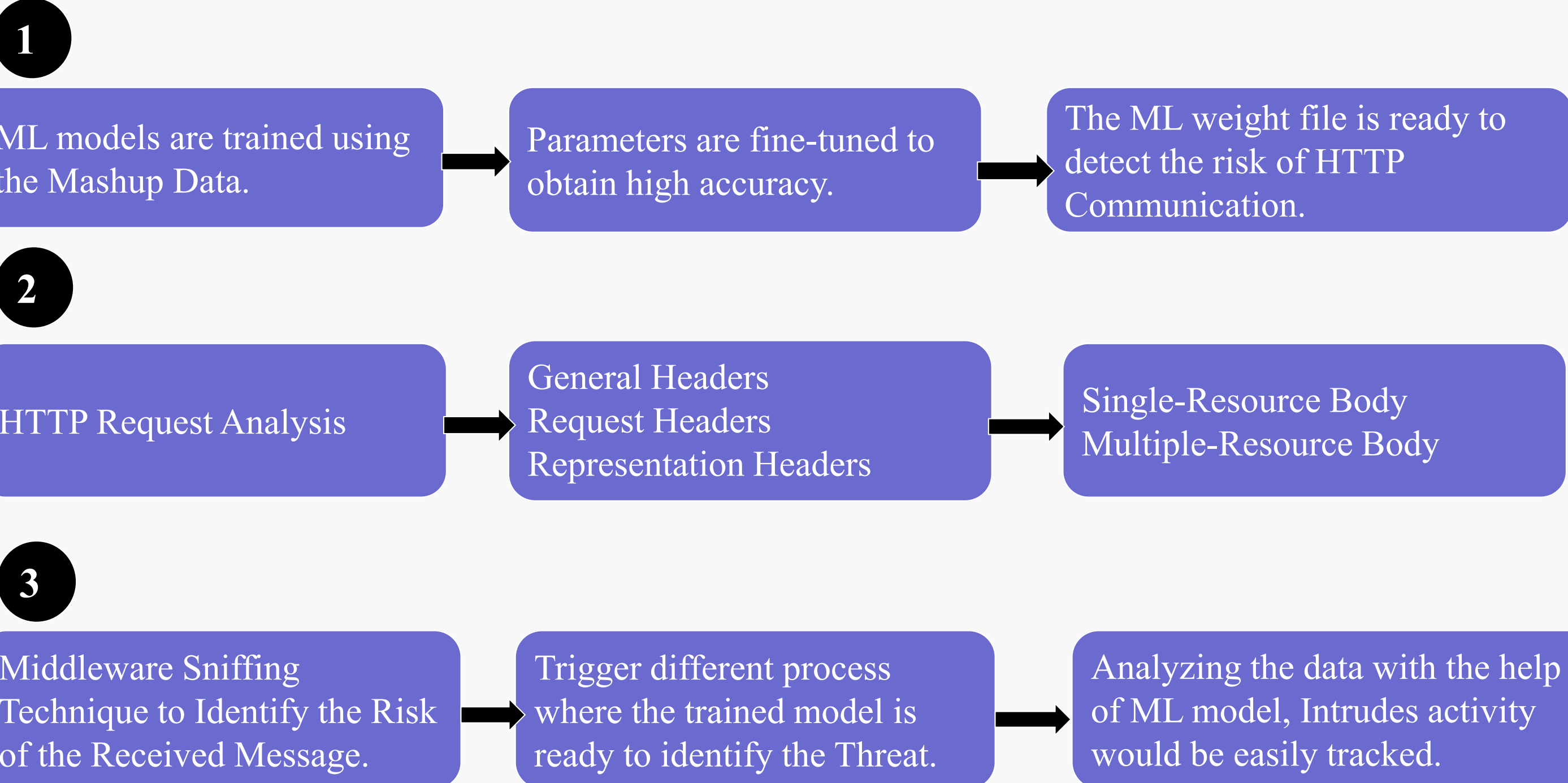
- There will always be new vulnerabilities, and new attack vectors introduced to evade the cyber detection method, In this context, **Mashup Data** with Machine Learning (ML) could be a new concept.
- Previous works focus less on industrial Red Teaming (**RT**) practice to design a cost-effective sustainable cyber echo system.
- A comprehensive analysis of ML models is necessary to determine which performs the best on the cyber attack signatures detection.
- A paradigm that could be utilized to make a usable **cyber protection process** will encourage our cyber warriors to adopt a re-engineering model.



Proposed Idea and Methodology

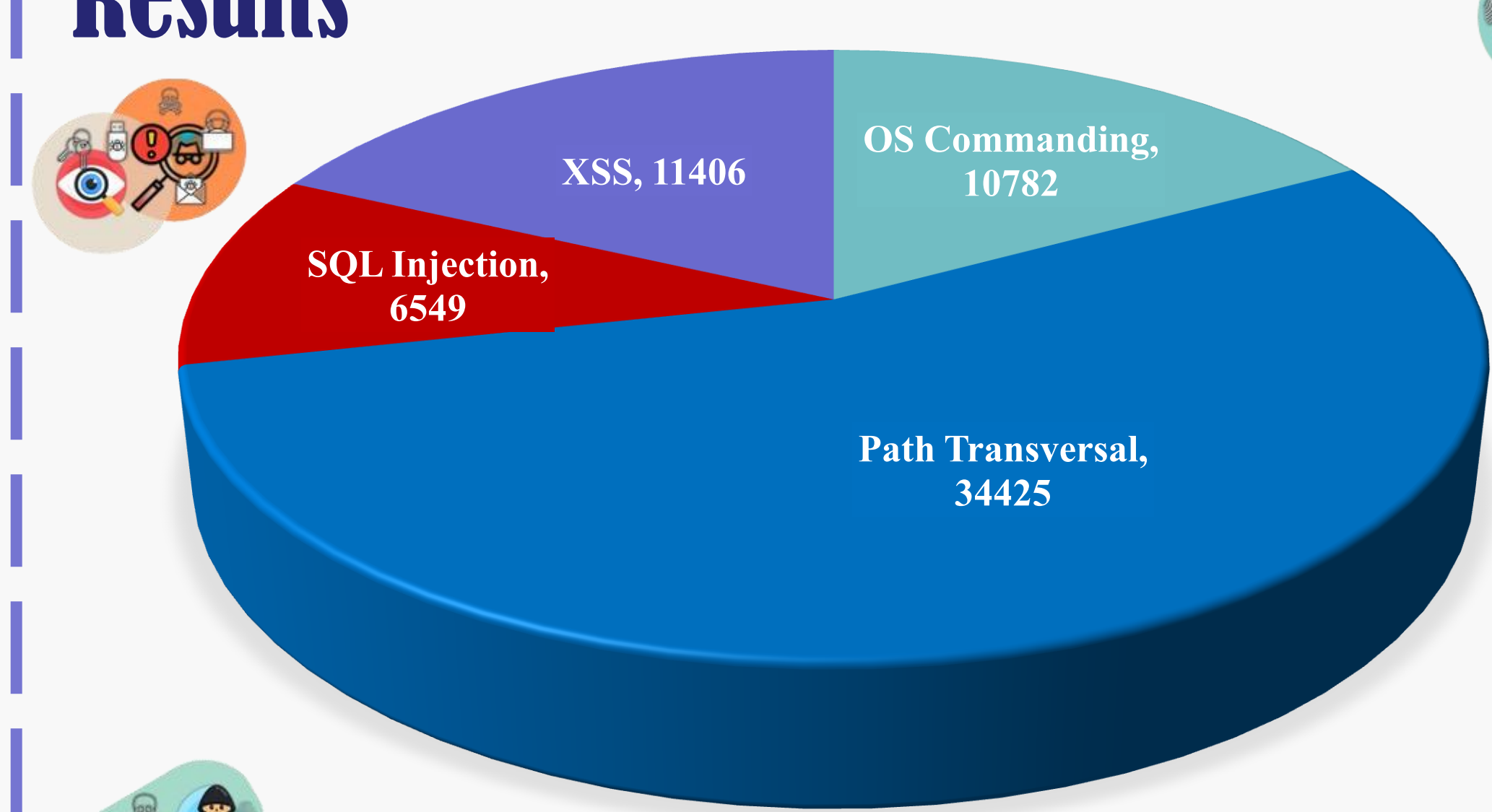


- Attacking Data annotation is a big challenge from the collected log of the DAST tools.
- There is no static method to find the dynamic attacking signatures.



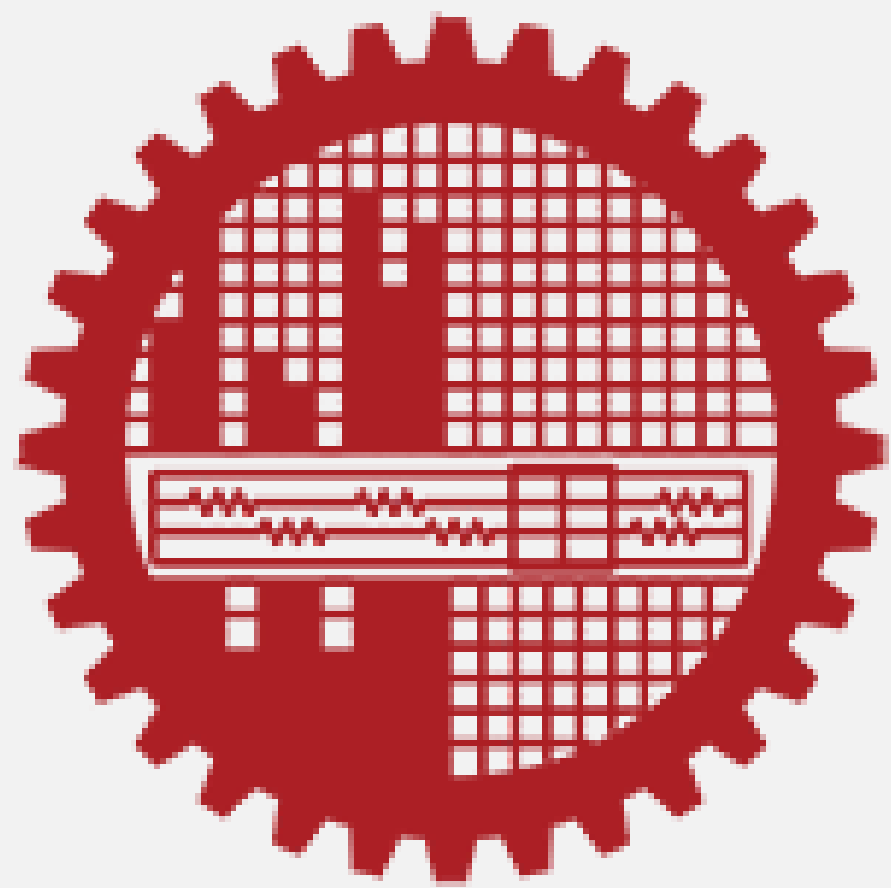
Results

COLLECTED ATTACKING SIGNATURES



ML Models	Accuracy(%)	Precession (%)	Recall(%)	F1 Score(%)
LSTM	98.04	98.10	98.04	98.06
SVM	98.13	98.14	98.13	98.11
CNN	98.27	98.00	98.00	98.00
RNN	98.14	98.21	98.14	98.16
Decision Tree	98.08	98.10	98.08	98.06
Logistic Regression	98.08	98.08	98.08	98.06
LSVM	98.17	98.26	98.17	98.19

A data-driven approach for forecasting of daily dengue cases.



Md. Reja E Rabbi Tonmoy
Supervisor: Dr. Md. Rubaiyat Hossain Mondal

Abstract

In order to address the crucial public health issue of dengue fever in tropical regions such as Bangladesh, our work introduces DTBoost, a cutting-edge machine learning algorithm intended for daily dengue incidence prediction. By integrating the extensive DengueWatchBD dataset (2016-2020), which includes detailed daily meteorological data, with the innovative I-FAT data-driven technique, DTBoost transcends traditional forecasting methods. It achieves a Mean Absolute Error (MAE) of 2.32 and an R^2 score of 0.96, showcasing its superior performance in predicting dengue outbreaks.

Background & Motivation

- The number of new malignant melanoma diagnosis has annually increased by 27 percent in the past decade.
- Climate variables are a major contributor to dengue cases like high temperature, heavy rainfall and humidity.



Figure 1: Meteorological Factors and Dengue Fever Incidence

- Many researchers use techniques like SVM, LSTM, RF, and others to experiment with monthly and weekly forecasts. However, dengue predictions did not achieve satisfactory accuracy.
- After addressing a significant gap in previous studies, we focused our investigation on daily forecasting using our proposed model, DTBoost.
- We created a new dataset called DengueWatchBD and a novel data-driven technique named I-FAT (Integration Framework with Aggregation Technique) to cope with limited data resources.

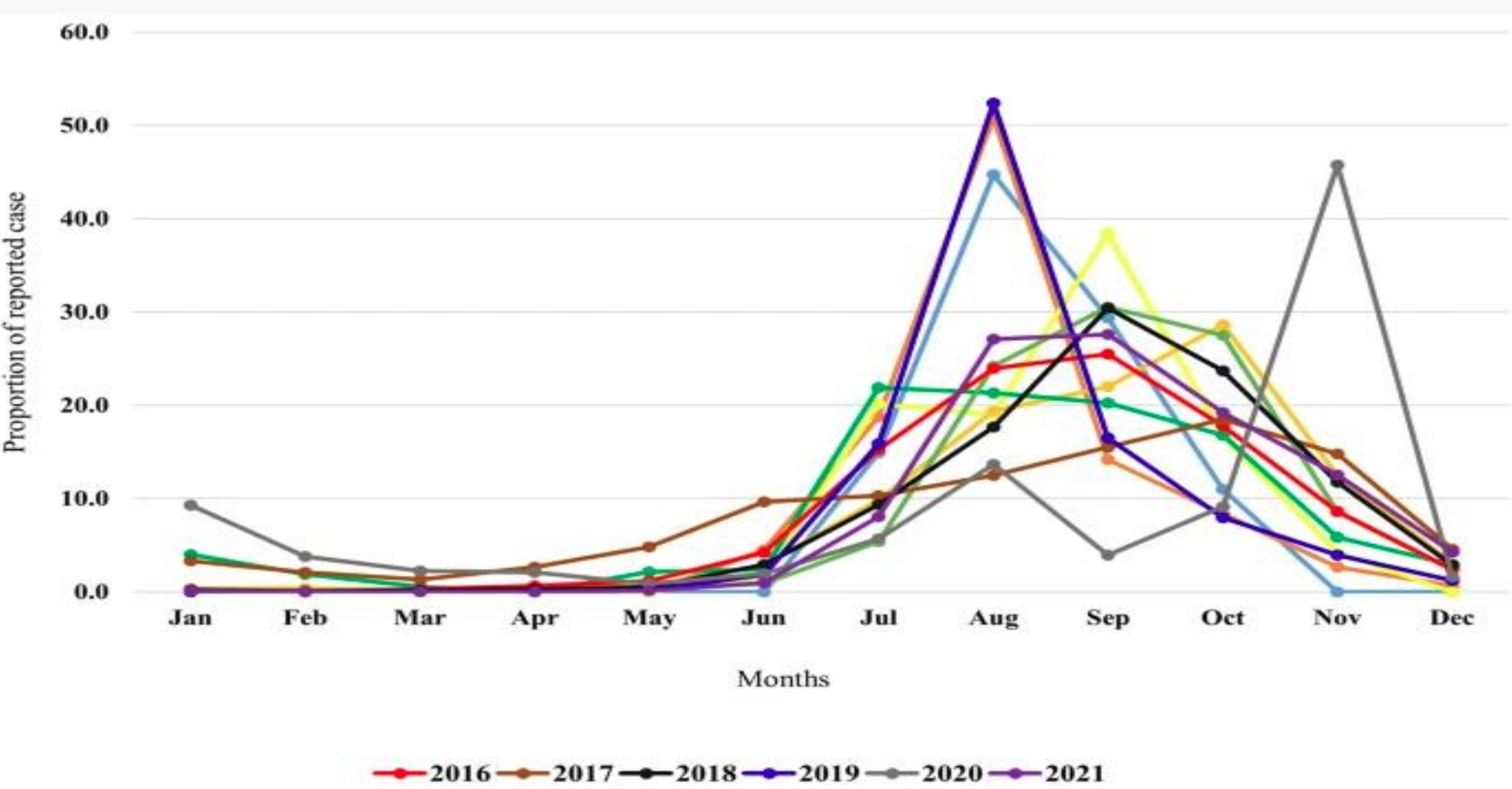


Figure 2: Seasonality of Dengue outbreaks in Bangladesh

Proposed Idea and Methodology

The simplified architecture of the proposed model is illustrated in Figure 3:

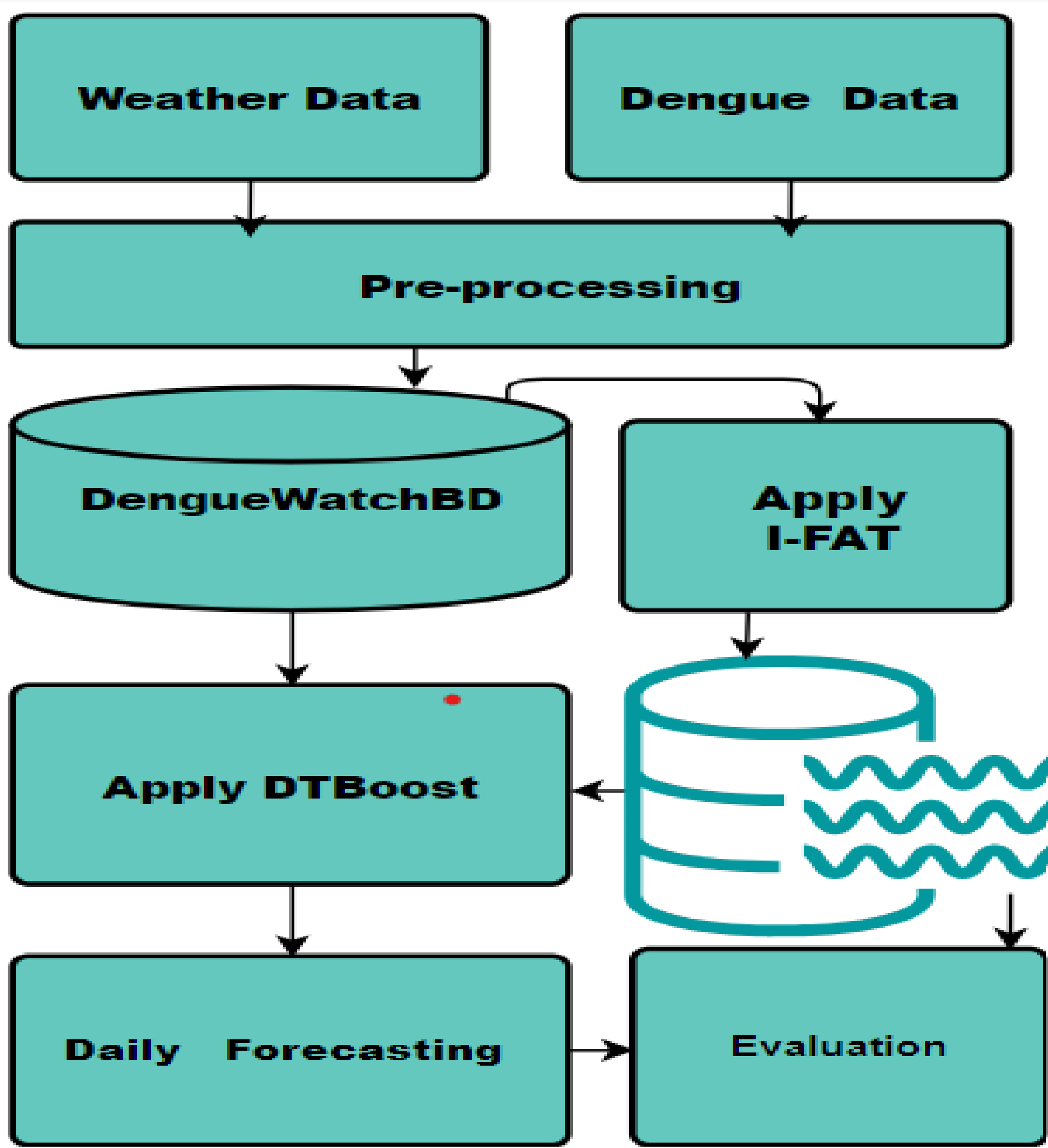


Figure 3: Simplified Conceptual Model of Proposed DTBoost Framework

We will collect the data from DGHS and BMD websites. After preprocessing the data, the I-FAT data-driven technique will be applied to the amalgamated data. Then we will apply our DTBoost model to the dataset. The model will be implemented using pandas, TensorFlow and Keras libraries of python.

Results

- ❖ We evaluated the performance of our daily dengue forecasting system on a test set of data. The train-test-split ratio is 0.2.
- ❖ The system achieved an accuracy of 96% and MAE of 2.32 in forecasting dengue cases on a daily basis.
- ❖ Our proposed DTBoost model outperformed other modes in terms of accuracy.
- ❖ We are working on to tune the models, and parameters and improve accuracy, sensitivity, and specificity.

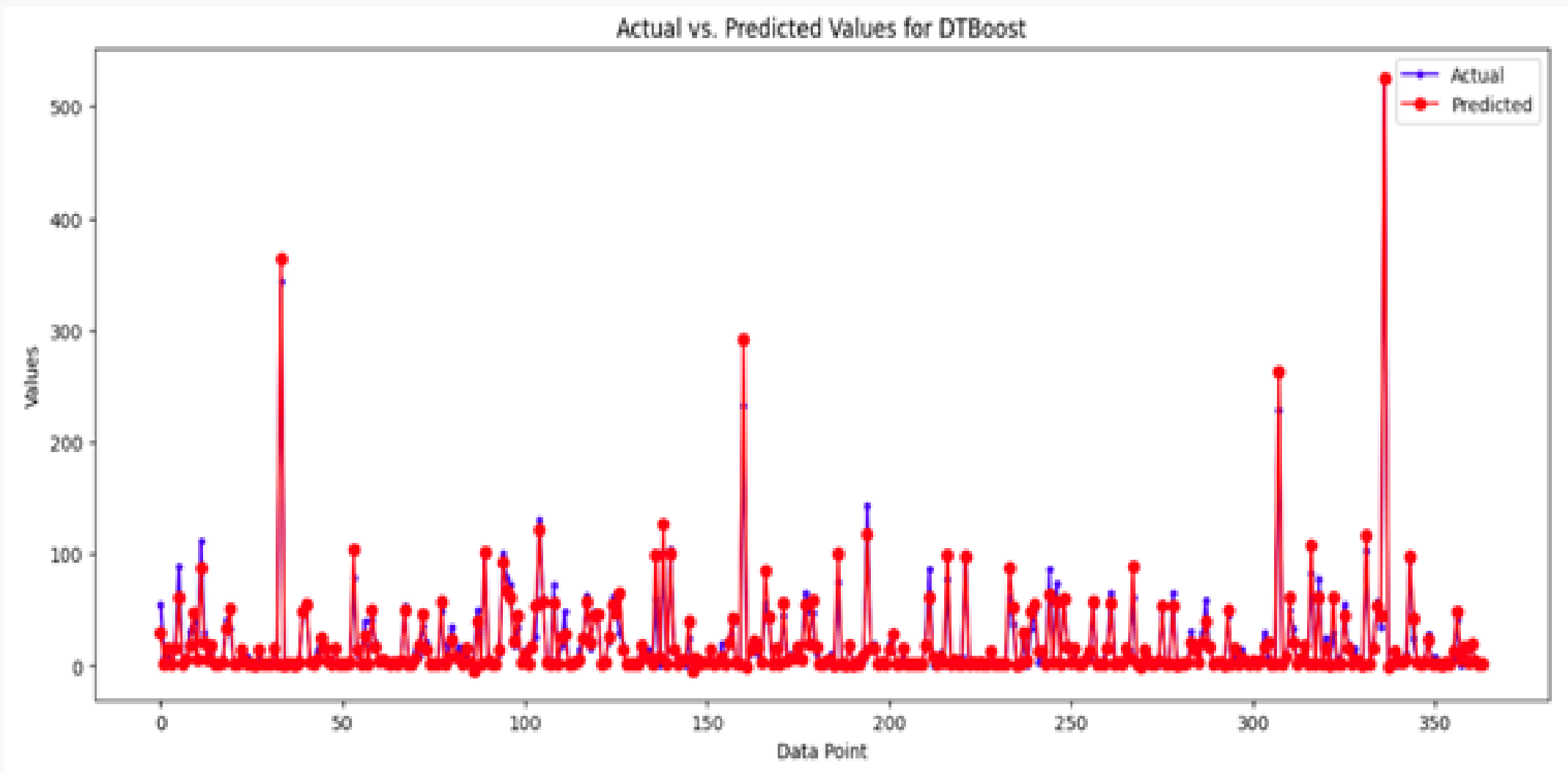
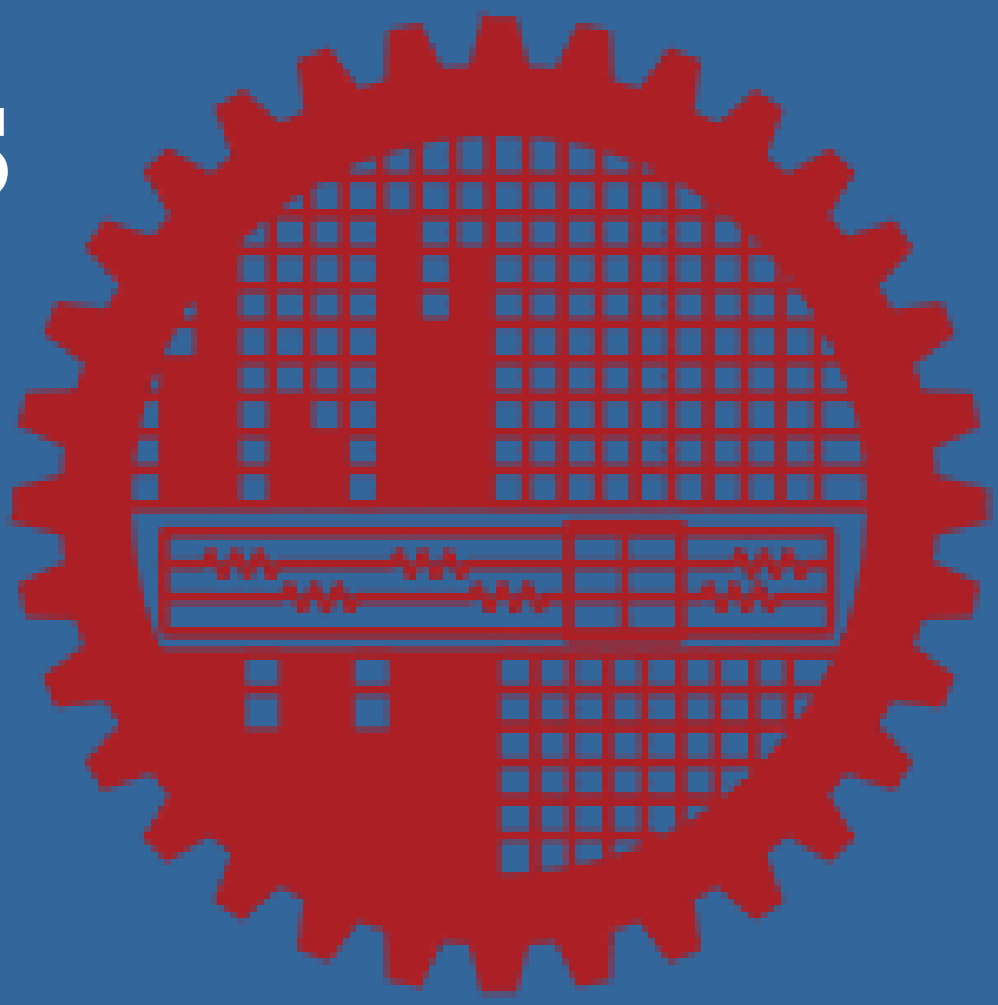


Figure 4: Actual and predicted Dengue Incidences with DTBoost

	Algorithm	Time Period	MAE	R2
0	Linear Regression	3 days	4.93	0.95
1	XGBoost	3 days	5.39	0.95
2	SVR	3 days	13.11	0.61
3	Linear Regression	1 week	6.48	0.78
4	XGBoost	1 week	4.93	0.92
5	SVR	1 week	8.14	0.69
6	Linear Regression	Fortnight	5.63	0.87
7	XGBoost	Fortnight	6.47	0.91
8	SVR	Fortnight	9.12	0.55
9	DTBoost (Proposed Model)	1 week	2.32	0.96

Figure 5: Comparative Analysis of Predictive Model's Performance

Exploration Of Pretrained Deep Learning Models for Wood Species Classification in Bangladeshi Forest with the Wood Tracker App.



Supervisor Name: Prof. Dr. Md. Saiful Islam
Name: Ashrafun Zannat

Abstract

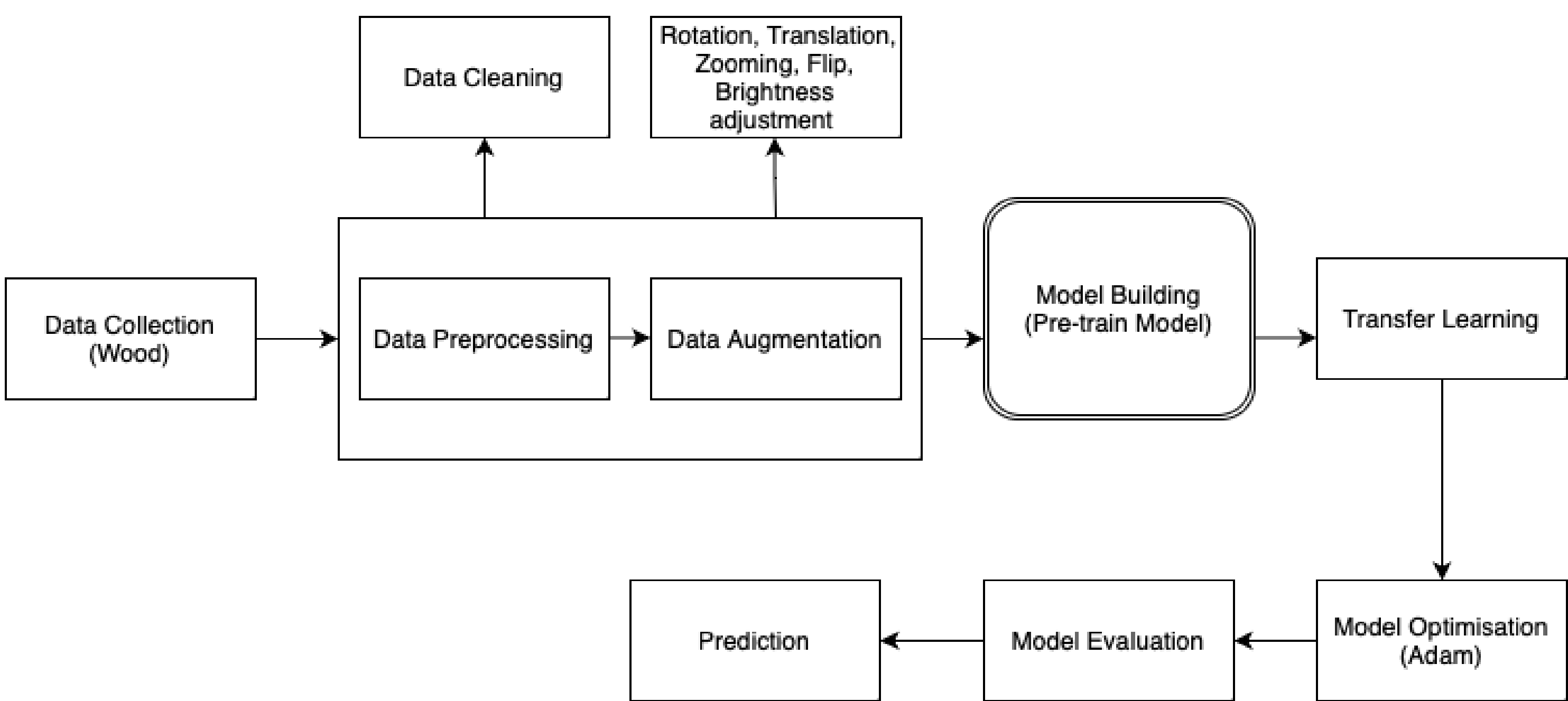
Mission of this work is to create a Bangladesh wood species dataset for meticulously created our own dataset, comprising over 1,800 high-quality images representing seven types of wood species indigenous to Bangladesh. To expand the dataset with data augmentation resulting approximately 7,125 images. Several deep learning models used for Wood Species Classification in Bangladeshi Forest with the Wood Tracker App. The deep learning model DenseNet121 was the clear winner in our thorough evaluation of seven pre-trained models, exhibiting unmatched quality with remarkable recall, precision, and an astounding F1 score of 0.9710, 0.9729, and 0.9717, respectively.

Background & Motivation

- In our industrial sectors, erroneous selection of wood species providing a vital role for destroying the forest.
- sustainable forest management is finding ways to continue to benefit from ecological services without compromising the forest's ability to provide those services.
- Forests currently absorb billions of tons of CO₂ globally every year. Mission of this work is to create a Bangladesh wood species dataset for selecting our required wood for our industries sectors and day to day life.
- Wood Species detection can raise awareness about the importance of plant biodiversity and conservation.
- Wood species detection can empower individuals and communities by providing accessible tools and knowledge about local flora.

Proposed Idea and Methodology

- Data Collection : Collect various data from the surrounding environment.
- Data Preprocessing : Clean, transform, and prepare data for analysis or modeling.
- Feature Selection : Selecting a subset of relevant features from a larger set of features that are available in a dataset.
- Model Development : Developing a machine learning model (CNN models like VGG16, ResNet-121, Xception, MobileNet).
- Model Evaluation: Model evaluation is the process of assessing the performance. Several commonly used evaluation metrics including accuracy, precision, recall, f1-score, ROC AUC.
- Model Optimizer : We compared different types of optimizer . Comparing between them, we select the best optimizer that provides the highest performance.



Data Set



Fig. 1. A few sample of Bangladeshi Wood

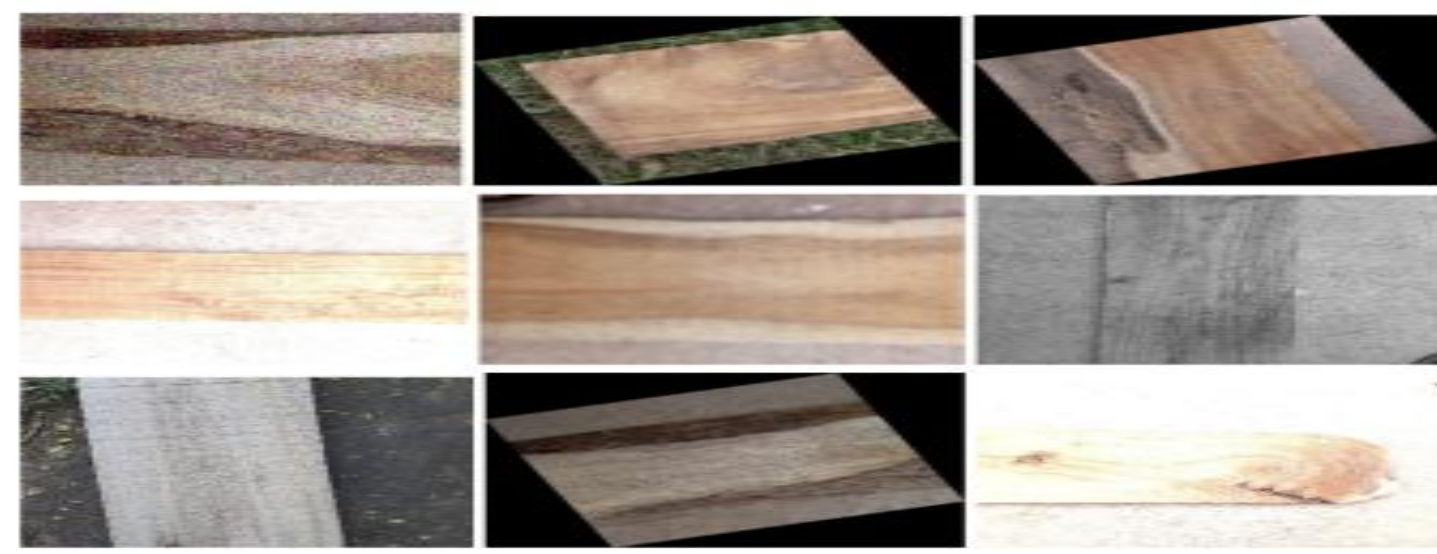


Fig. 2. Sample image after applying data augmentation.

Predicted Classification

Deep learning models classifies 7 classes. The Classes are Eucalyptus, Loha , Akashmoni, Mehogoni ,Minjiri, Raintree ,Kathal.



Result

pre-trained models were tested individually utilizing transfer learning techniques on our dataset during the experiment, and the results of training and validation were carefully recorded. DenseNet121 performed brilliantly in the transfer learning scenario as well , with training accuracy of 97.09% and validation accuracy of 97.14% .With a recall of 97.10%,precision of 97.29%, and an F1 score of 97.17%, DenseNet121.

Table 1:Comparative result Analysis (Train accuracy and Validation accuracy)

Model Name	Train Accuracy	Validation Accuracy
DenseNet121	97.09%	97.14%
Xception	97.00%	93.64%
ResNet152V2	96.82%	95.36%
InceptionV3	95.70%	93.93%

Chart 1:Train Accuracy and Validation Accuracy

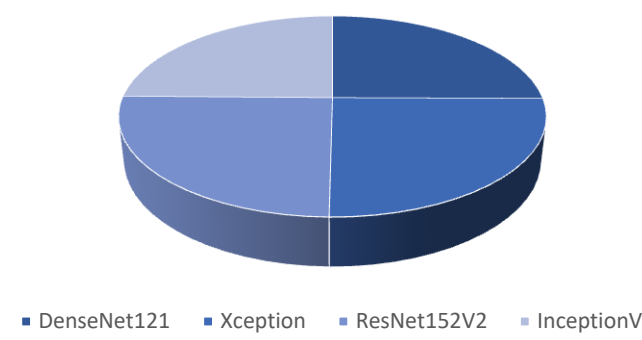
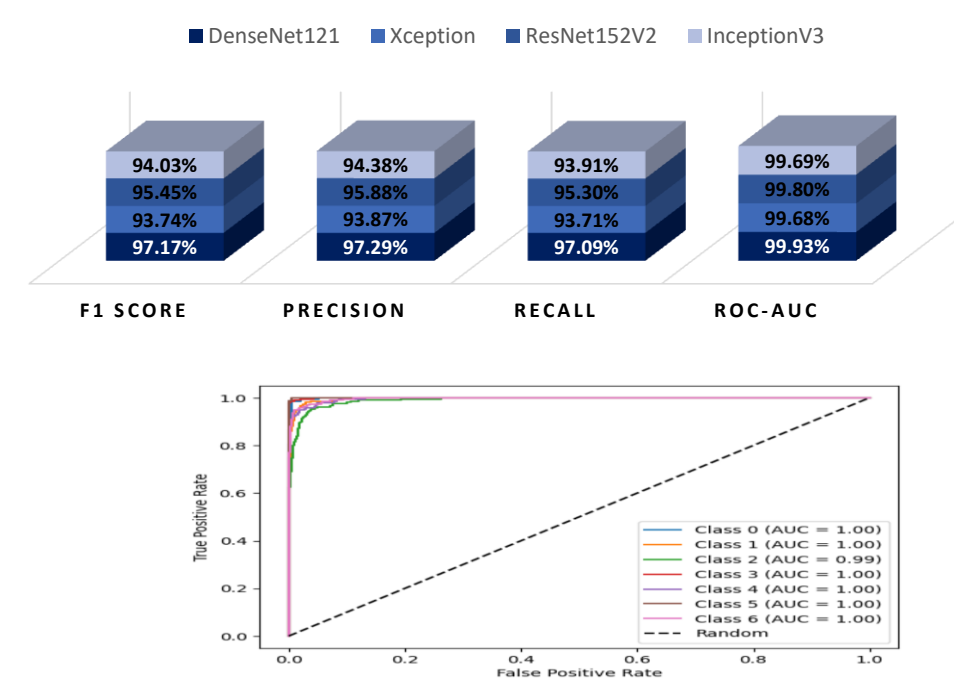


Table 2:Comparative result Analysis (F1,Precision,Recall,ROC-AUC)

Model Name	F1 Score	Precision	Recall	ROC-AUC
DenseNet121	97.17%	97.29%	97.09%	99.93%
Xception	93.74%	93.87%	93.71%	99.68%
ResNet152V2	95.45%	95.88%	95.30%	99.80%
InceptionV3	94.03%	94.38%	93.91%	99.69%

CHART: 2



Discussion

We propose a wood tracker app that will be very helpful for selecting necessary wood species for our industrial sectors. In the realm of deep learning and artificial intelligent systems, the analysis of wood species identification holds immense importance. The identification of wood species is a critical aspect of forestry, conservation, and various industrial applications. Different wood species possess distinct characteristics that determine their suitability for specific purposes, such as construction.

Conclusion

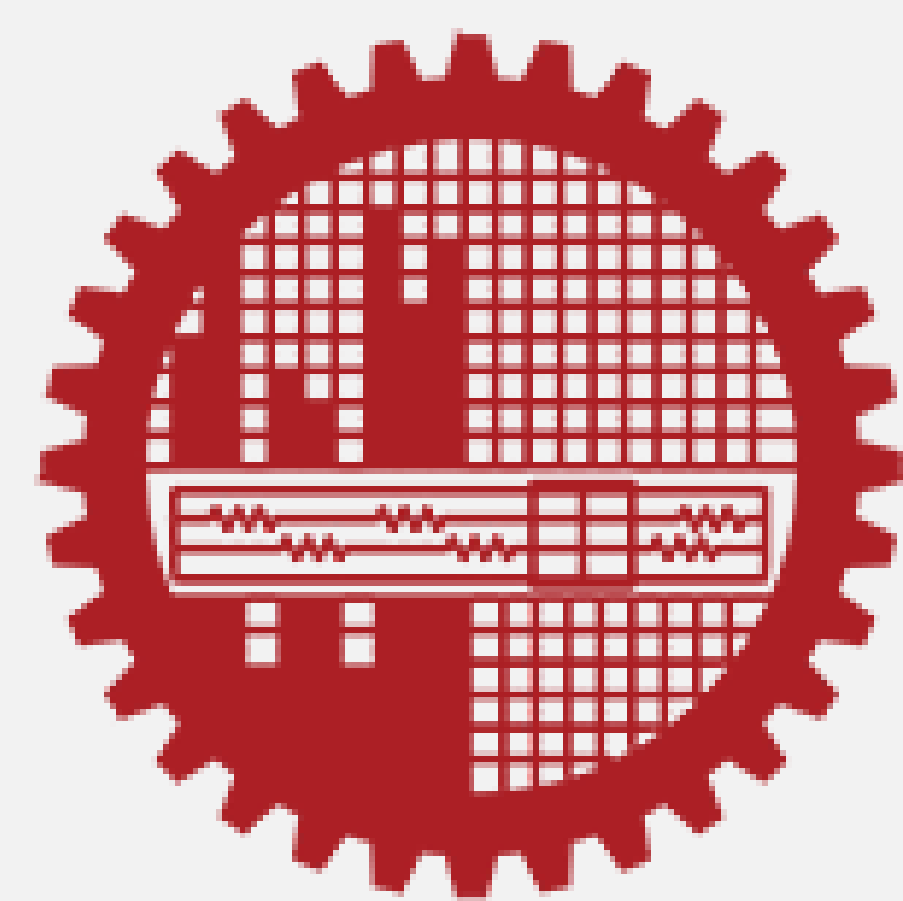
- Due to cutting of forests and Pollution, lot of medicinal Plants leaves have almost become extinct. So, there is an immediate need for us to identify them and replant them for the use of next generations.
- Wood species detection can aid botanists and researchers in studying and documenting various plant species

References

- [1] Chollet, Franc,ois. "Xception: Deep learning with depthwise separable convolutions." Proceedings of the IEEE conference on computer vision and pattern recognition. 2017.
- [2] He, Kaiming, et al. "Identity mappings in deep residual networks." Computer Vision–ECCV 2016: 14th European Conference, Amsterdam, The Netherlands, October 11–14, 2016, Proceedings, Part IV 14. Springer International Publishing. 2016.
- [3] Szegedy, Christian, et al. "Rethinking the inception architecture for computer vision." Proceedings of the IEEE conference on computer vision and pattern recognition. 2016.
- [4] Simonyan, Karen, and Andrew Zisserman. "Very deep convolutional networks for large-scale image recognition." arXiv preprint arXiv:1409.1556 (2014).
- [5] Simonyan, Karen, and Andrew Zisserman. "Very deep convolutional networks for large-scale image recognition." arXiv preprint arXiv:1409.1556 (2014).

DEVELOPMENT OF A SYSTEM FOR ASSESSMENT OF THE NEURODEVELOPMENTAL IMPAIRMENTS (NDI) OF CHILD

Farhad Rakib and Hossen A Mustafa



Abstract

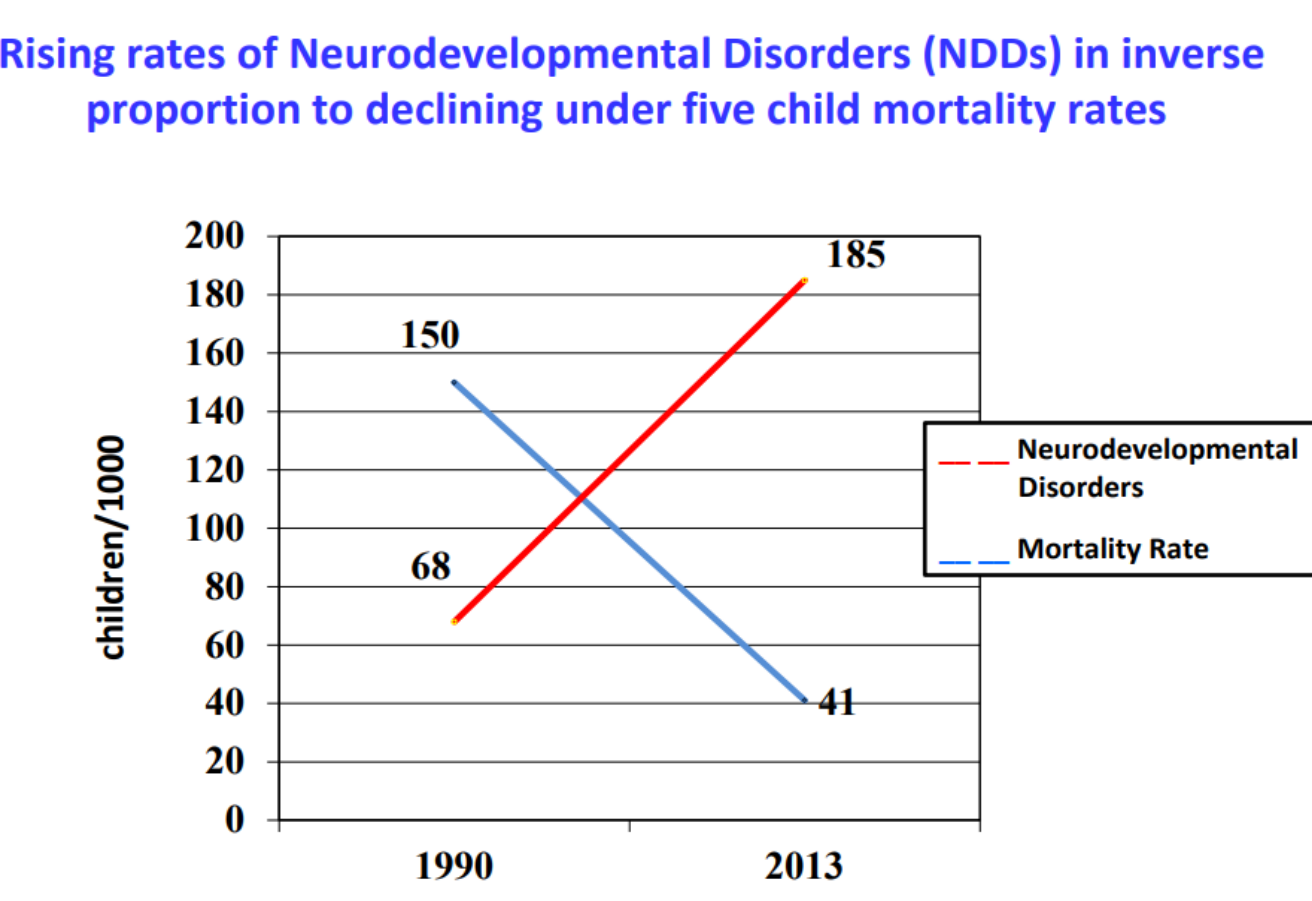
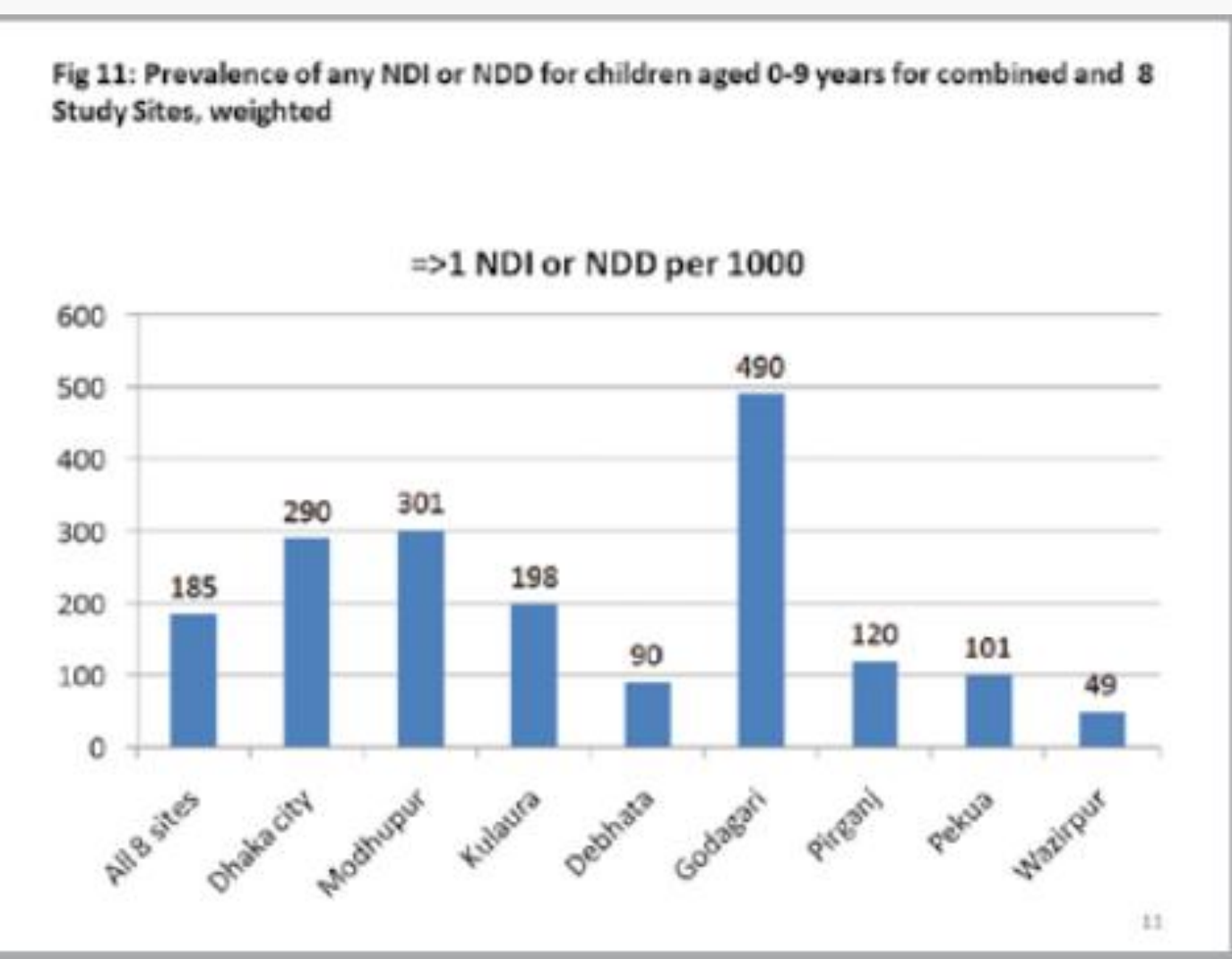
Neuro-developmental impairments (NDI) encompass a range of conditions that affect the growth and development of the brain. Severity of NDIs can be reduced by early detection. To detect NDIs, a manual assessment tool named RAPID NUERO-DEVELOPMENTAL ASSESSMENT (RNDA) is developed that evaluates neuro-developmental impairments across multiple domains, such as hearing, cognitive abilities, and vision. Our work proposes to eliminate the manual process by developing an algorithm and based on the algorithm, a web based system which detects the NDI level by using the same methodology of RNDA. Another target of this work is to create a dataset of different NDI levels and apply different machine learning model to detect the NDIs in an efficient and trustworthy manner .

Background

- ❑ In 2021, Early Diagnostics and Early Intervention in Neurodevelopmental Disorders—Age Dependent Challenges and Opportunities.
- ❑ Again in 2021, Development of a Pathway for Multidisciplinary Neurodevelopmental Assessment and Diagnosis in Children and Young People.
- ❑ In 2016, Smart autism – a mobile, interactive and integrated framework for screening and confirmation of autism.
- ❑ Again in 2016 , Smart-NDA: A cloud based framework for smart device integrated automated NDI screening tool.

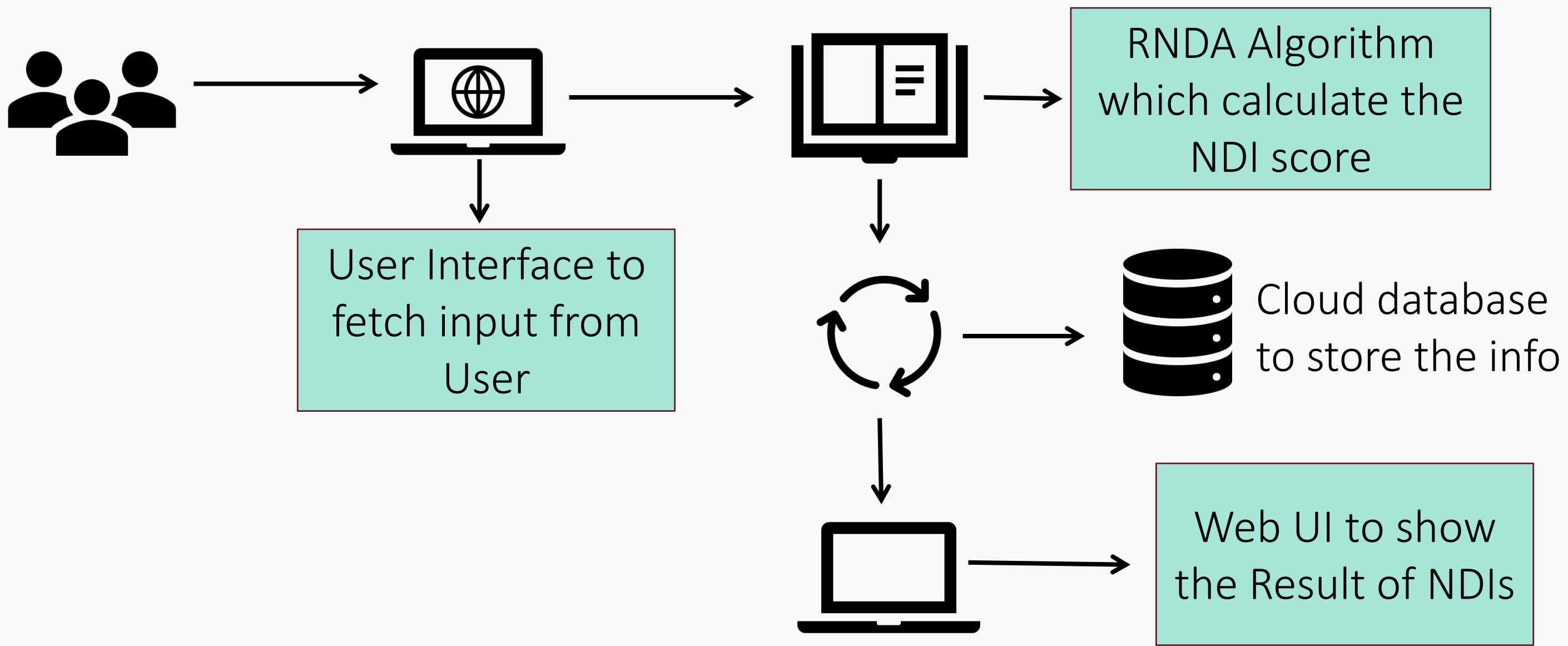
Motivation

- ❑ In recent years , NDI among the children of 0-9 years old are increasing at an alarming rate.
- ❑ Previous works are based on theoretical approach and no analytical implementations are done.
- ❑ A comprehensive procedure or algorithm is to be developed which can be used to determine the NDI in future.

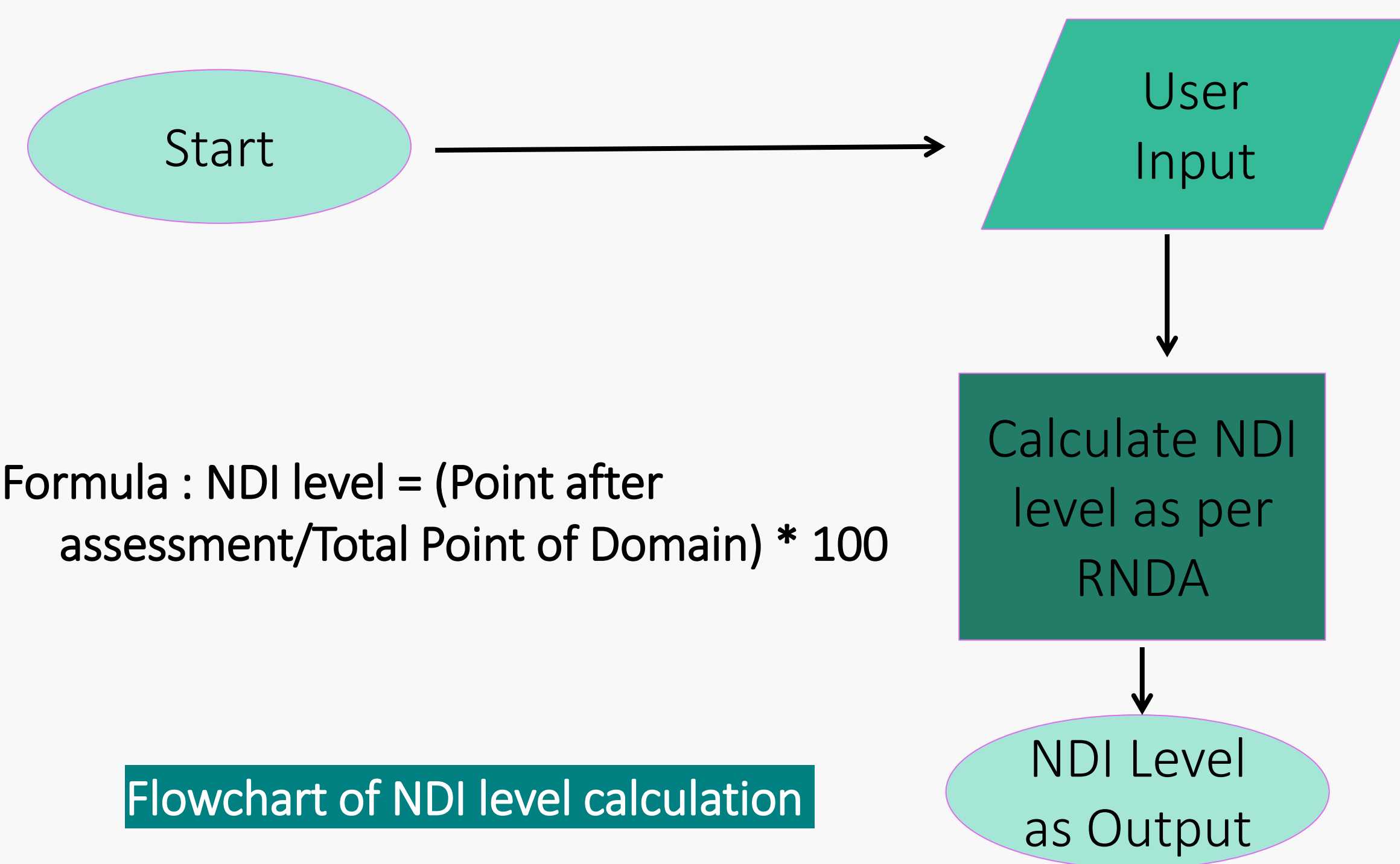


Proposed Idea and Methodology

- ❑ To eliminate the manual (paper based approach), we are proposing a web based interactive system to determine the level of NDIs by using RNDA



- ❑ To calculate the level of NDIs an algorithm will be developed by using the calculative approach defined by RNDA and a ML based analysis will be done.



Results

Patient Information Form

SI No. Name

Address

Diagnosis Date of Admission

Date of Discharge

```
def GrossMotorNDIlevelcalculation(age, assessmentTotal):
    """
    Calculate Gross Motor NDI (Neurological Developmental Index) level based
    on age and assessmentTotal.

    Parameters:
    - age: Age of the individual.
    - assessmentTotal: Numeric value obtained from a gross motor assessment.

    Returns:
    - NDI_level: Gross Motor NDI level.
    """
    NDI_level = (assessmentTotal / 16) * 100
    return NDI_level
```

Result of NDI's using RNDA

Your NDI level of Growth Motor is Normal

Your NDI level of Vision is Moderate. Please take RNDA again after 3 months.