ROBUST SPAMMER DETECTION USING COLLABORATIVE NEURAL NETWORK IN INTERNET OF THINGS APPLICATION

Project Members:

- Biradar Nitesh (207R1A05M4)
- Polaboina Sreeja (207R1A05M6)
- Marelli Ramya (207R1A05M1)

CONTENTS :

Abstract :

- In this paper, we propose a novel approach leveraging collaborative neural networks (CNNs) for robust spammer detection. By harnessing the power of collaborative learning, our model effectively captures intricate patterns and anomalies present in IoT data streams.
- We design a CNN architecture capable of learning from multiple sources of data generated by IoT devices distributed across diverse networks. The collaborative framework enables the model to generalize well across different IoT environments, enhancing its robustness against adversarial attacks and evolving spamming techniques.

Existing System:

- Traditional spam detection methods often rely on rule-based approaches or machine learning algorithms trained on centralized data, which may not be suitable for the distributed and decentralized nature of IoT networks.
- Rule-based spam detection systems, while simple and easy to implement, often lack adaptability to evolving spamming techniques and struggle to generalize across diverse IoT deployments. These systems typically rely on predefined rules or thresholds to flag suspicious activities, making them prone to false positives and negatives in dynamic IoT environments.

Disadvantages of Existing System:

- Limited Adaptability
- High False Positive Rates
- Vulnerability to Adversarial Attacks
- Scalability Issues
- Complexity and Maintenance

Proposed System:

- In this paper, a Collaborative neural network based Spammer detection mechanism (Co-Spam) is proposed to solve the above problems. Co-Spam combines both semantic and behavioral patterns to solve spammer detection problems.
- In our work, the speech contents and behavior records of users at different time stamps are first viewed as their feature sequences. At each timestamp, a bidirectional auto recorder (Bi-AE) is developed to model semantic characteristics, and graph convolutional network (GCN) is designed to learn the embedding of behavior patterns

Advantages of Proposed System:

- Distributed Processing
- Adaptability
- Resource Efficiency
- Scalability
- Security and Privacy
- Real-World Validation
- Collaborative Learning

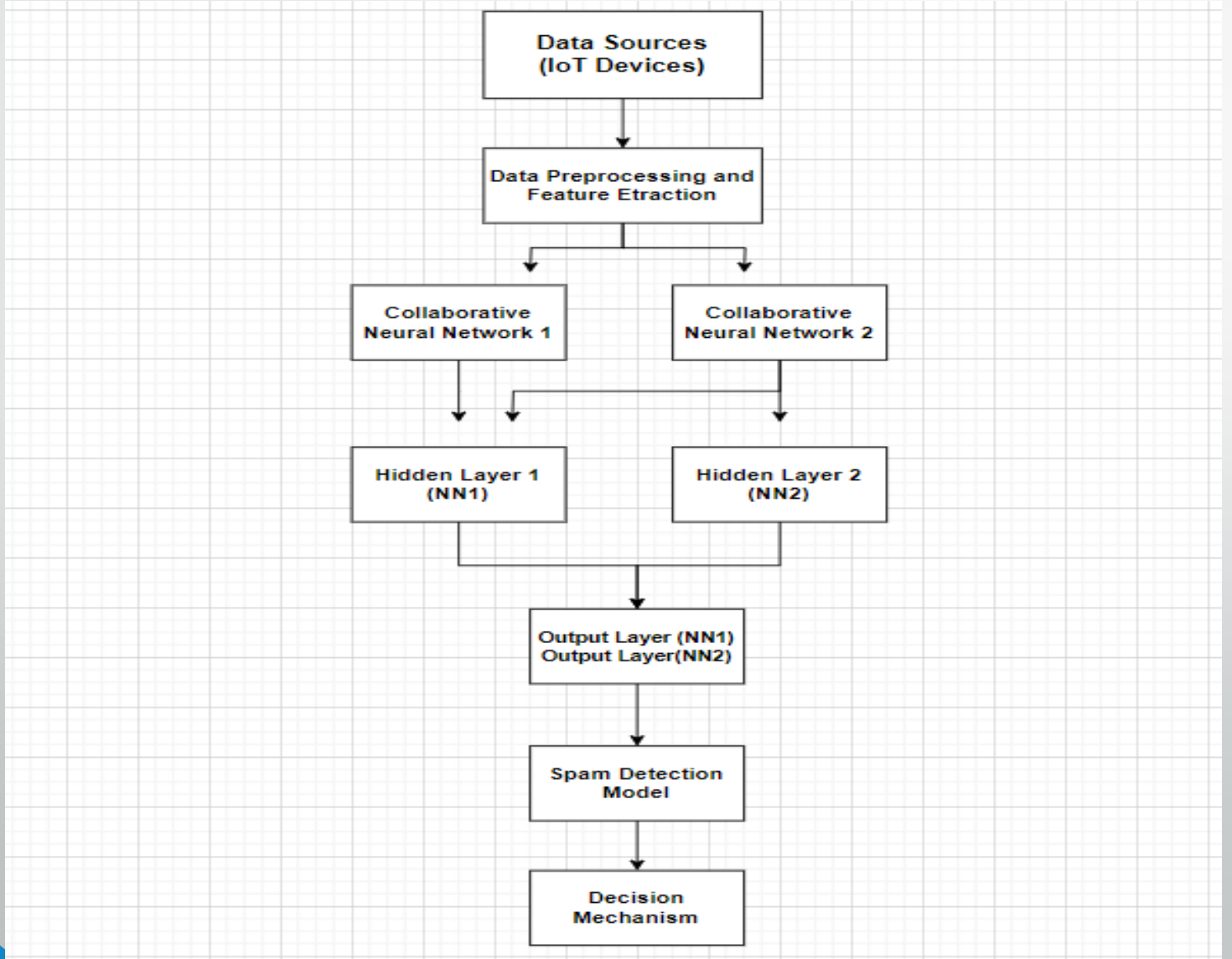
Hardware Requirements:

- Processor : Pentium-IV
- RAM : 4GB
- Hard Disk : 20 GB
- Key Board : Standard Windows Keyboard
- Mouse : 2 or 3 Button Mouse
- Monitor : SVGA

Novelty of the Project:

At its core, the project proposes a collaborative neural network (CNN) model specifically tailored for the distributed and diverse nature of IoT environments. Unlike traditional centralized approaches, this model is deployed directly on edge devices, forming a collaborative network where lightweight spam detection agents collaborate to learn and adapt to local data streams. This edge-centric architecture minimizes reliance on centralized processing, reducing communication overhead and enhancing scalability and efficiency, particularly in resource-constrained IoT deployments.

Architecture :

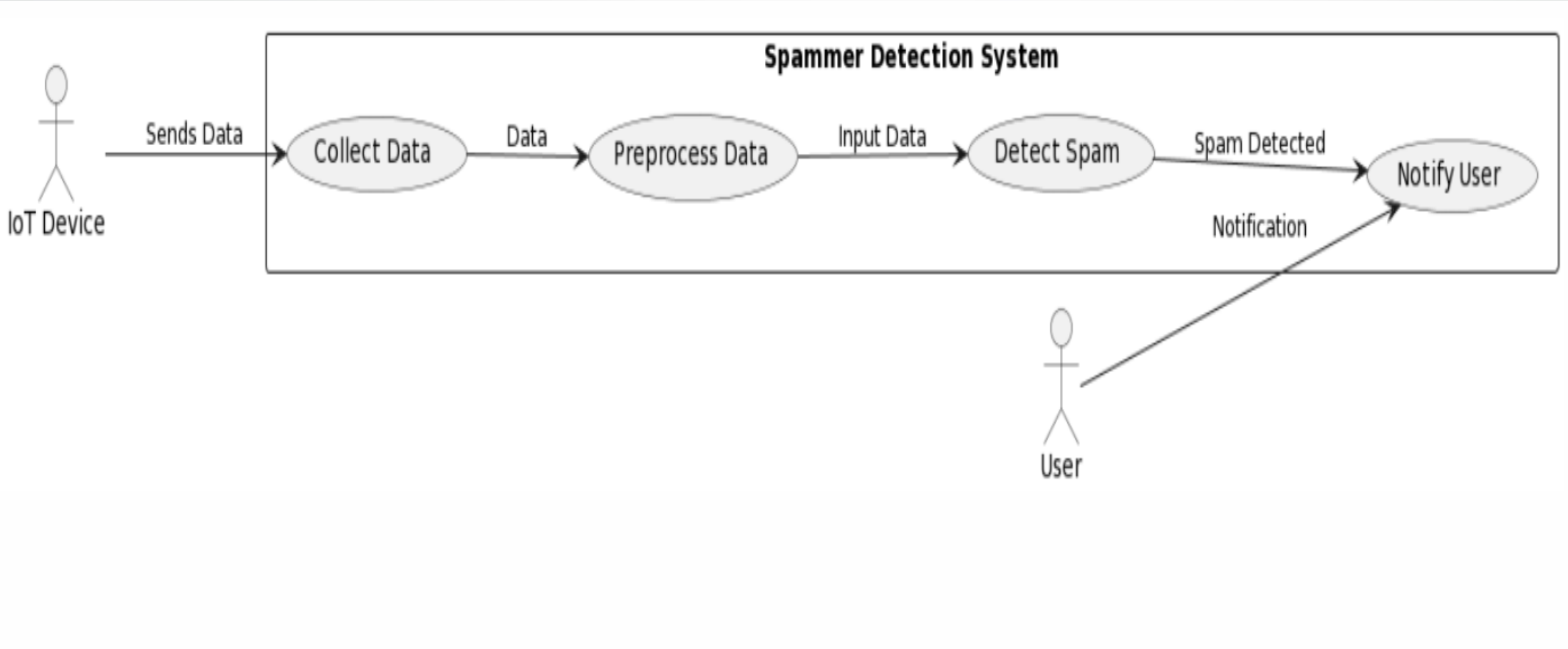


Modules :

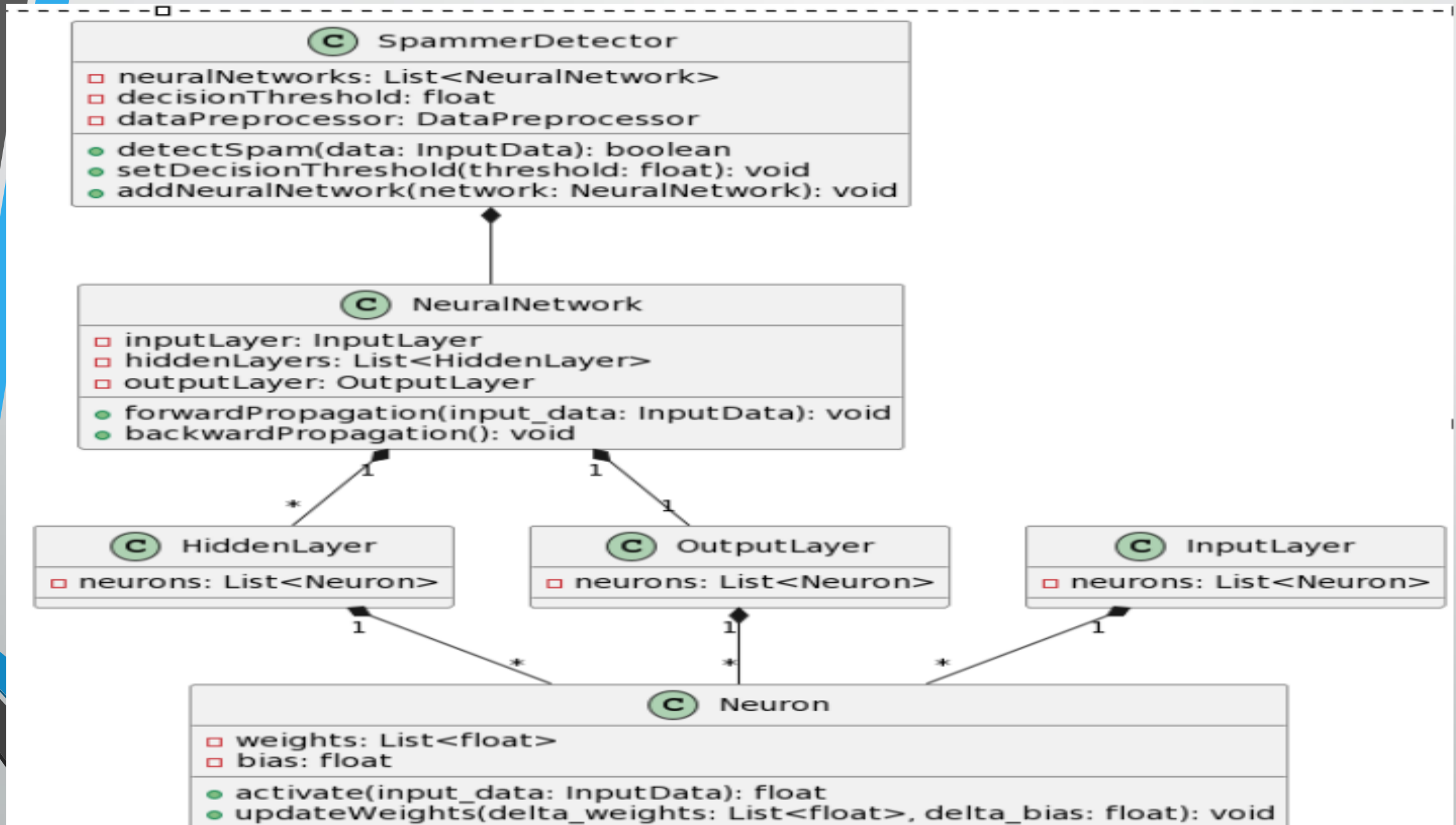
- Data Preprocessing
- Feature Extraction
- Collaborative Neural Network

UML Diagram :

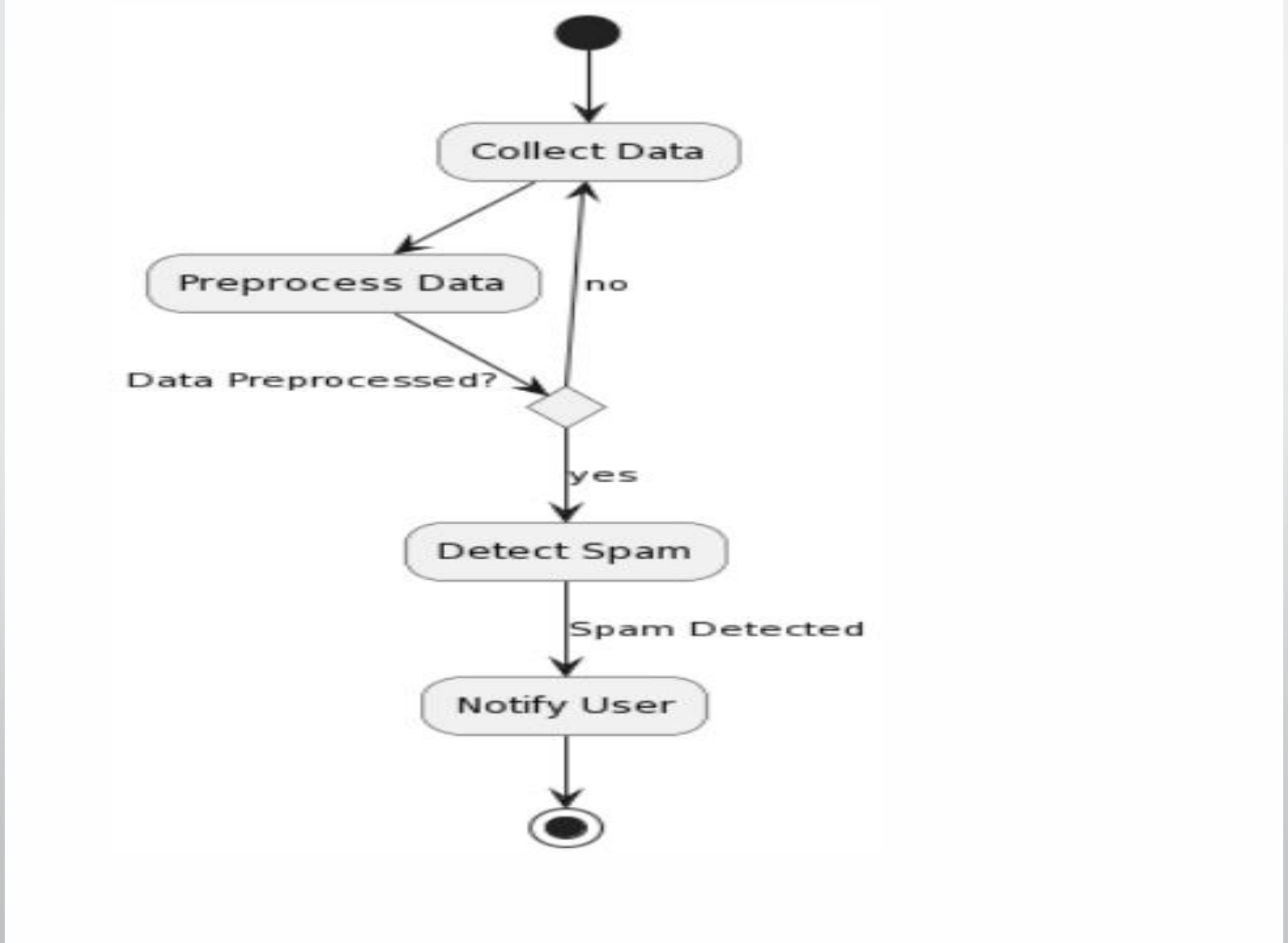
- **Use Case Diagram:**



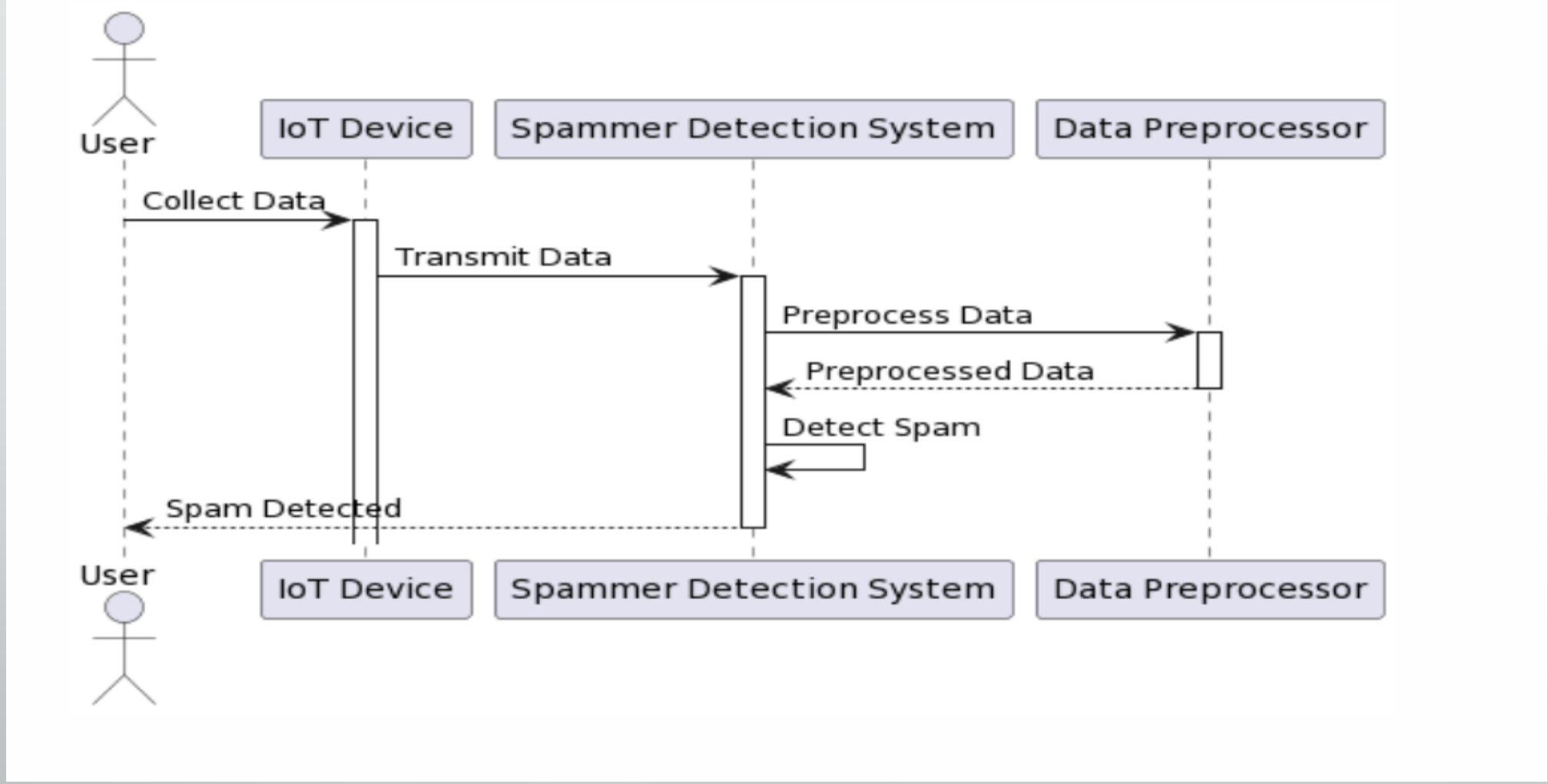
- **Class Diagram:**



- **Activity Diagram:**



- **Sequence Diagram :**



Sample Code:

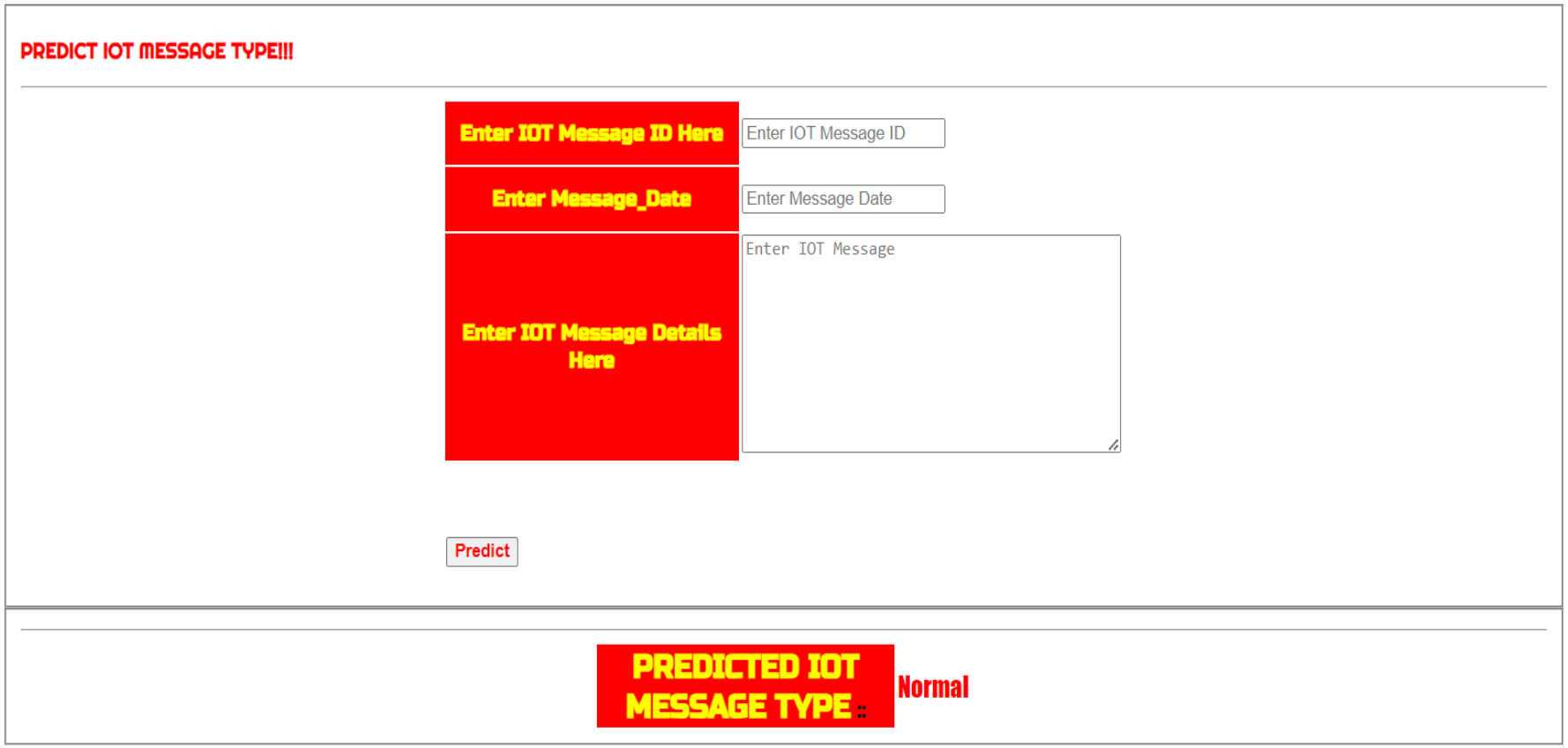
```
from django.shortcuts import render
from django.template import RequestContext
from django.contrib import messages
import pymysql
from django.http import HttpResponse
from django.conf import settings
from django.core.files.storage import FileSystemStorage
import matplotlib.pyplot as plt
import re
import cv2
import numpy as np
from string import punctuation
from nltk.corpus import stopwords
import nltk
from nltk.stem import WordNetLemmatizer
from nltk.stem import PorterStemmer
import os
from nltk.tokenize import word_tokenize
```

```
stop_words = set(stopwords.words('english'))
lemmatizer  = WordNetLemmatizer()
porter = PorterStemmer()


def LCS(l1,l2):  #LCS method
    s1 = word_tokenize(l1)
    s2 = word_tokenize(l2)
    dp = [[None]*(len(s1)+1)  for i in range(len(s2)+1)]
    for i in range(len(s2)+1):
        for j in range(len(s1)+1):
            if i == 0 or j == 0:
                dp[i][j] = 0
            elif s2[i-1] == s1[j-1]:
                dp[i][j] = dp[i-1][j-1]+1
            else:
                dp[i][j] = max(dp[i-1][j] , dp[i][j-1])
    return dp[len(s2)][len(s1)]


def cleanPost(doc):
    tokens = doc.split()
    table = str.maketrans('', '', punctuation)
    tokens = [w.translate(table) for w in tokens]
```

Results :


PREDICT IOT MESSAGE TYPE!!!
Enter IOT Message ID Here
Enter IOT Message ID
Enter Message_Date
Enter Message Date
Enter IOT Message Details Here
Enter IOT Message
Predict
PREDICTED IOT MESSAGE TYPE ::
Normal

An Efficient Spam Detection Technique for IoT Devices Using Machine Learning

Browse IOT Data Sets and Train & Test | View Trained and Tested Accuracy in Bar Chart | View Trained and Tested Accuracy Results | View Prediction Of IOT Message Type | View IOT Message Type Ratio | Download IOT Message Predicted Data Sets | View IOT Message Type Ratio Results | View All Remote Users | Logout

VIEW ALL REMOTE USERS !!!

USER NAME	EMAIL	Mob No	Country	State	City
Rajesh	Rajesh123@gmail.com	9535866270	India	Karnataka	Bangalore
Manjunath	tmksmanju13@gmail.com	9535866270	India	Karntaka	Bangalore
hp	hp@gmail.com	9090909090	india	telangana	hyderabad
hp	hp@gmail.com	2345678193	india	telangana	Hyderabad
sreeja	sree@gmail.com	1234567809	india	telangana	Hyderabad
hp	hp@gmail.com	8888888888	india	telangana	hyderabad
then	then@gmail.com	123456789	india	telangana	MEDAK

An Efficient Spam Detection Technique for IoT Devices Using Machine Learning

Browse IOT Data Sets and Train & Test | View Trained and Tested Accuracy in Bar Chart | View Trained and Tested Accuracy Results | View Prediction Of IOT Message Type | View IOT Message Type Ratio

Download IOT Message Predicted Data Sets | View IOT Message Type Ratio Results | View All Remote Users | Logout

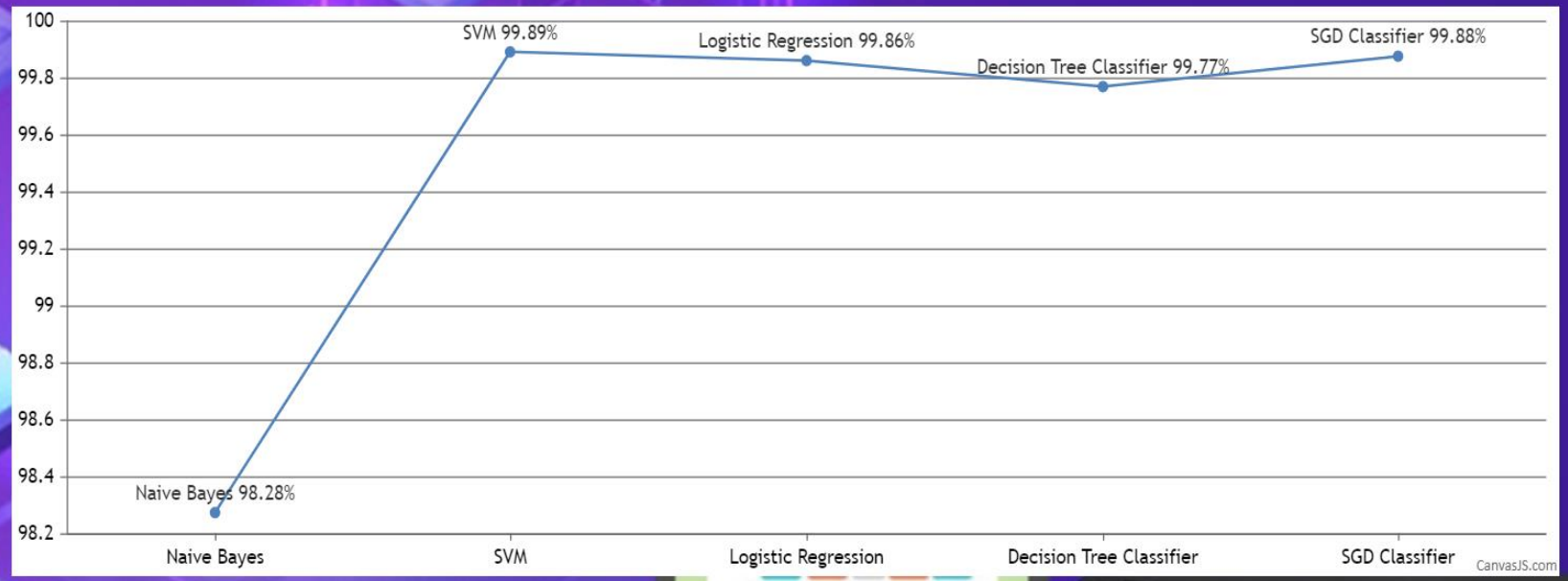
An Efficient Spam Detection Technique for IoT Devices Using Machine Learning

Browse IOT Data Sets and Train & Test | View Trained and Tested Accuracy in Bar Chart | View Trained and Tested Accuracy Results | View Prediction Of IOT Message Type | View IOT Message Type Ratio

Download IOT Message Predicted Data Sets | View IOT Message Type Ratio Results | View All Remote Users | Logout

PIE CHAERT

LINE CHART



View IOT Message Prediction Type Details !!!

IOT Message Id	IOT Message Date	IOT Message	Prediction
4	meter 7268 nov allocation	kimberly vaughn / hou / ect on 12 / 10 / 99 01 : 52 pm - lauri a allen 12 / 09 / 99 01 : 20 pm to : kimberly vaughn / hou / ect @ ect , anita luong / hou / ect @ ect cc : howard b camp / hou / ect @ ect , mary m	Normal
18367	start date : 2 / 6 / 02 ; hourahead hour : 24 ;	log messages : parsing file - - > > o : \ portland \ westdesk \ california scheduling \ iso final schedules \ 2002020624 . txt ! ! ! general sql error . couldn ' t update ; currently locked by user ' admin ' on machine ' nahou - trdts 5 ' . table - - - - energy import /	Spam
18367	start date : 2 / 6 / 02 ; hourahead hour : 24 ;	log messages : parsing file - - > > o : \ portland \ westdesk \ california scheduling \ iso final schedules \ 2002020624 . txt ! ! ! general sql error . couldn ' t update ; currently locked by user	Spam

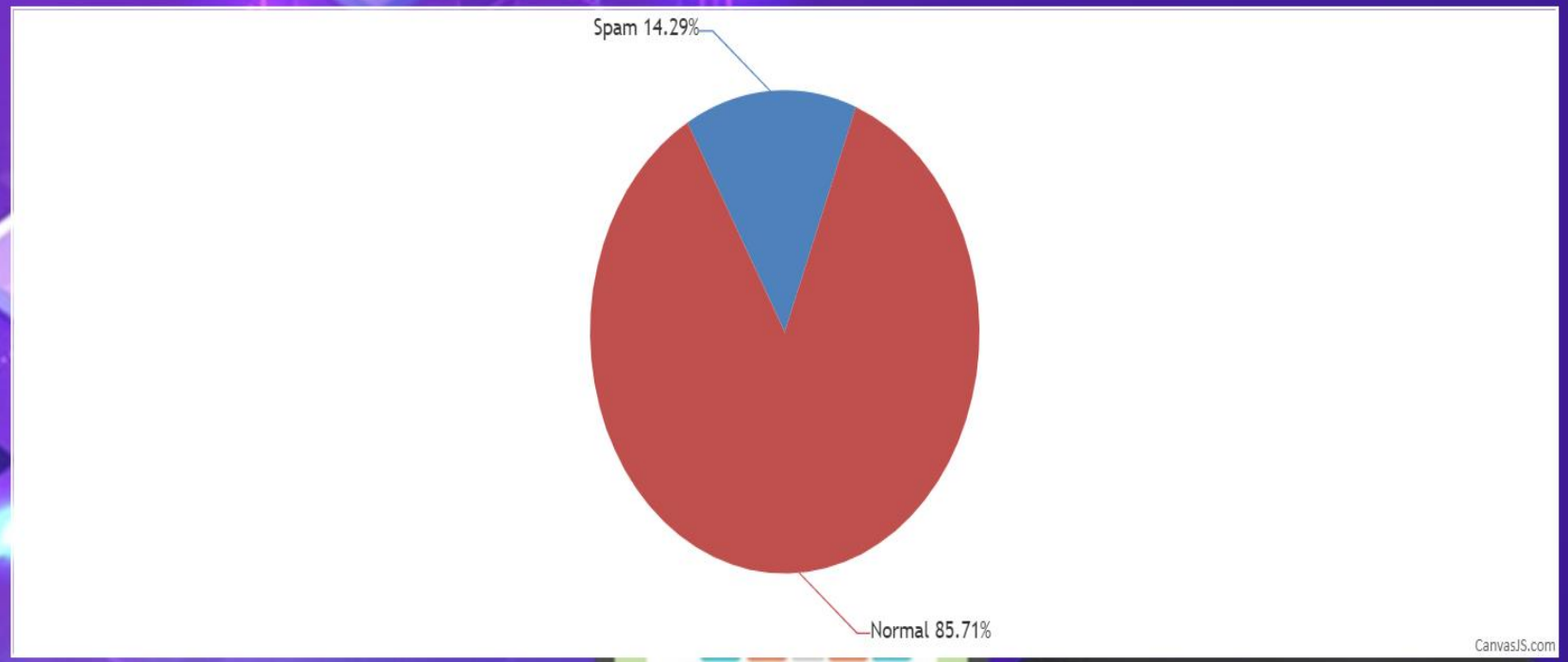
An Efficient Spam Detection Technique for IoT Devices Using Machine Learning

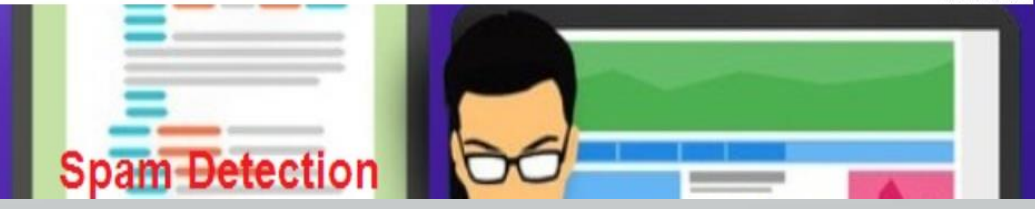
Browse IOT Data Sets and Train & Test | View Trained and Tested Accuracy in Bar Chart | View Trained and Tested Accuracy Results | View Prediction Of IOT Message Type | View IOT Message Type Ratio

Download IOT Message Predicted Data Sets | View IOT Message Type Ratio Results | View All Remote Users | Logout

PIE CHAERT

LINE CHART





Conclusion:

In conclusion, the project represents a significant advancement in the field of spam detection within Internet of Things (IoT) applications. By proposing a collaborative neural network (CNN) architecture deployed at the edge, the project introduces a novel approach that addresses the unique challenges of spam detection in distributed and dynamic IoT environments. Through extensive experimentation and evaluation on real-world datasets, the effectiveness and reliability of the proposed system have been demonstrated across diverse domains.

Future Scope :

- Multi-Modal Data Fusion
- Enhanced Collaborative Learning
- Real-Time Response Mechanisms
- Privacy-Preserving Techniques
- User Education and Awareness

References :

- https://www.researchgate.net/publication/342325957_Robust_Spammer_Detection_Using_Collaborative_Neural_Network_in_Internet_of_Thing_Applications
- https://waseda.elsevierpure.com/en/publications/robust-spammer-detection-using-collaborative-neural-network-in-in
- https://researchonline.federation.edu.au/vital/access/manager/Repository/vital:16538;jsessionid=FFE4AC966CDB5C2CD9BDF681546E8610?view=null&f0=sm_creator%3A%22Imran%2C+Muhammad%22&f1=sm_creator%3A%22Bashir%2C+Ali%22&sort=null&f2=sm_subject%3A%224009+Electronics%2C+Sensors+and+Digital+Hardware%22

Github Link:

- https://github.com/sreeja2702/major-project

Thank You.