

A
Mini Project
On
Use Of Artificial Neural Networks to Identify Fake Profiles
(Submitted in partial fulfillment of the requirements for the award of Degree)

BACHELOR OF TECHNOLOGY
IN
COMPUTER SCIENCE AND ENGINEERING

By
KHAJA MUBASHIRUDDIN (207R1A05L6)
P. BHARAT SIMHA REDDY (207R1A05N5)
GADDE SAYI KHUSHHAL (207R1A05L0)

UNDER THE GUIDANCE OF
S. APARNA
(Assistant Professor)



DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING
CMR TECHNICAL CAMPUS
UGC AUTONOMOUS

(Accredited by NAAC, NBA, Permanently Affiliated to JNTUH, Approved by AICTE, New Delhi)
Recognized Under Section 2(f) & 12(B) of the UGC Act, 1956, Kandlakoya (V), Medchal Road,
Hyderabad-501401.

2020-2024

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING



CERTIFICATE

This is to certify that the project entitled **“USE OF ARTIFICIAL NEURAL NETWORKS TO IDENTIFY FAKE PROFILES”** being submitted by **KHAJA MUBASHIRUDDIN (207R1A05L6), P. BHARAT SIMHA REDDY (207R1A05N5) & G. SAYI KHUSHHAL (207R1A05L0)** in partial fulfilment of the requirements for the award of the degree of B.Tech in Computer Science and Engineering to the Jawaharlal Nehru Technological University Hyderabad, is a record of bonafide work carried out by them under our guidance and supervision during the year 2023-2024.

The results embodied in this thesis have not been submitted to any other University or Institute for the award of any degree or diploma.

S. APARNA
(Associate Professor)
INTERNAL GUIDE

DR. A. RAJI REDDY
DIRECTOR

DR. K. SRUJAN RAJU
HOD

EXTERNAL EXAMINER

Submitted for viva voice Examination held on_____

ACKNOWLEDGEMENT

Apart from the efforts of us, the success of any project depends largely on the encouragement and guidelines of many others. We take this opportunity to express our gratitude to the people who have been instrumental in the successful completion of this project.

We take this opportunity to express my profound gratitude and deep regard to my guide **S. Aparna**, for her exemplary guidance, monitoring and constant encouragement throughout the project work. The blessing, help and guidance given by her shall carry us a long way in the journey of life on which we are about to embark.

We also take this opportunity to express a deep sense of gratitude to the Project Review Committee (PRC) **G.Vinsh Shanker, Dr. J. Narasimharao, Ms. Shilpa, & Dr. K. Maheswari** for their cordial support, valuable information and guidance, which helped us in completing this task through various stages.

We are also thankful to **Dr. K. Srujan Raju**, Head, Department of Computer Science and Engineering for providing encouragement and support for completing this project successfully.

We are obliged to **Dr. A. Raji Reddy**, Director for being cooperative throughout the course of this project. We also express our sincere gratitude to Sri. **Ch. Gopal Reddy**, Chairman for providing excellent infrastructure and a nice atmosphere throughout the course of this project.

The guidance and support received from all the members of **CMR Technical Campus** who contributed to the completion of the project. We are grateful for their constant support and help.

Finally, we would like to take this opportunity to thank our family for their constant encouragement, without which this assignment would not be completed. We sincerely acknowledge and thank all those who gave support directly and indirectly in the completion of this project.

KHAJA MUBASHIRUDDIN	(207R1A05L6)
P. BHARAT SIMHA REDDY	(207R1A05N5)
GADDE SAYI KHUSHHAL	(207R1A05L0)

ABSTRACT

The pervasive presence of fake profiles on social networking platforms, such as Facebook, poses significant challenges in ensuring user security and trust. This project endeavors to address this issue by employing the power of machine learning, specifically artificial neural networks (ANNs), to determine the authenticity of incoming Facebook friend requests. The project begins with the collection of comprehensive datasets comprising Facebook profiles, encompassing diverse attributes related to user information, friend networks, and online activity. These datasets serve as the foundation for training and validating an ANN-based model designed for binary classification, distinguishing between authentic and fake friend requests. Several essential components are discussed within the project, including the selection of pertinent classes and libraries such as TensorFlow or PyTorch for ANN implementation, data preprocessing, and feature engineering. Feature selection involves a careful consideration of parameters extracted from social network profiles, including profile completeness, image quality, friend connections, and activity frequency.

Central to the ANN's operation is the sigmoid function, which produces probability scores between 0 and 1, serving as a measure of the likelihood of a friend request being authentic. The project delves into the workings of the sigmoid function in the context of binary classification, elucidating its role in rendering meaningful predictions. Moreover, the project elucidates the intricate process of weight determination and adaptation during ANN training. These weights represent the neural network's learned parameters and are fine-tuned iteratively through optimization algorithms such as gradient descent, ensuring the model's proficiency in authenticating friend requests.

LIST OF FIGURES/TABLES

FIGURE NO	FIGURE NAME	PAGE NO
Figure 3.1	Project Architecture for Use of Artificial Neural Networks to Identify Fake Profiles	7
Figure 3.2	Use Case Diagram for Use of Artificial Neural Networks to Identify Fake Profiles	8
Figure 3.3	Class Diagram for Use of Artificial Neural Networks to identify Fake Profiles	9
Figure 3.4	Sequence Diagram for Use of Artificial Neural Networks to Identify Fake Profiles	10
Figure 3.5	Activity Diagram for Use of Artificial Neural Networks to Identify Fake Profiles	11

LIST OF SCREENSHOTS

SCREENSHOT NO.	SCREENSHOT NAME	PAGE NO.
Screenshot 5.1	Upload Social Media Dataset	16
Screenshot 5.2	Select Dataset	16
Screenshot 5.3	Preprocessing the Dataset	17
Screenshot 5.4	Run ANN Algorithm	17
Screenshot 5.5	Observe Accuracy	18
Screenshot 5.6	Generate ANN Accuracy and Loss Graph	18
Screenshot 5.7	ANN Accuracy and Loss Graph	19
Screenshot 5.8	Load Test Data	19
Screenshot 5.9	Fake/Genuine Profile Detection	20

TABLE OF CONTENTS

ABSTRACT	i
LIST OF FIGURES	ii
LIST OF SCREENSHOTS	iii
1. INTRODUCTION	1
1.1 PROJECT SCOPE	1
1.2 PROJECT PURPOSE	1
1.3 PROJECT FEATURES	1
2. SYSTEM ANALYSIS	2
2.1 PROBLEM DEFINITION	2
2.2 EXISTING SYSTEM	2
2.2.1 LIMITATIONS OF THE EXISTING SYSTEM	3
2.3 PROPOSED SYSTEM	3
2.3.1 ADVANTAGES OF PROPOSED SYSTEM	3
2.4 FEASIBILITY STUDY	4
2.4.1 ECONOMIC FEASIBILITY	4
2.4.2 TECHNICAL FEASIBILITY	4
2.4.3 SOCIAL FEASIBILITY	5
2.5 HARDWARE & SOFTWARE REQUIREMENTS	5
2.5.1 HARDWARE REQUIREMENTS	5
2.5.2 SOFTWARE REQUIREMENTS	6
3. ARCHITECTURE	7
3.1 PROJECT ARCHITECTURE	7
3.2 DESCRIPTION	7
3.3 USE CASE DIAGRAM	8
3.4 CLASS DIAGRAM	9
3.5 SEQUENCE DIAGRAM	10
3.6 ACTIVITY DIAGRAM	11

TABLE OF CONTENTS

4. IMPLEMENTATION	12
4.1 SAMPLE CODE	12
5. SCREENSHOTS	16
6. TESTING	21
6.1 INTRODUCTION TO TESTING	21
6.2 TYPES OF TESTING	21
6.2.1 UNIT TESTING	21
6.2.2 INTEGRATION TESTING	21
6.2.3 FUNCTIONAL TESTING	22
6.3 TEST CASES	23
6.3.1 CLASSIFICATION	23
7. CONCLUSION & FUTURE SCOPE	24
7.1 PROJECT CONCLUSION	24
7.2 FUTURE SCOPE	24
8. BIBLIOGRAPHY	25
8.1 REFERENCES	25
8.2 GITHUB LINK	25

1. INTRODUCTION

1. INTRODUCTION

1.1 PROJECT SCOPE

This project is titled “Use of Artificial Neural Networks to identify Fake Profiles”. In this project, our scope encompasses a comprehensive exploration of the challenges posed by fake profiles on social media platforms. To begin, we will meticulously collect extensive datasets from various social media platforms, ensuring representation from both genuine and fake profiles. These datasets will form the basis of our research, allowing us to develop and fine-tune an artificial neural network (ANN) that can effectively discern authenticity from fraudulent profiles.

1.2 PROJECT PURPOSE

The overarching purpose of our project is to confront and mitigate the pervasive issue of fake profiles on social media platforms. Central to our endeavor is the mission to enhance the security and trustworthiness of online interactions, fostering a safer digital environment for all users. One of our primary objectives is to bolster online security. By leveraging artificial neural networks and machine learning, our project aims to provide a potent tool for detecting and mitigating the risks associated with fake profiles. We intend to empower users with the means to make informed decisions and safeguard their personal information from potential threats.

1.3 PROJECT FEATURES

This project boasts a comprehensive set of features, including diverse data collection, a robust artificial neural network (ANN) architecture capable of handling multi-modal data, meticulous data preprocessing, feature engineering to select pertinent attributes, extensive model training and optimization, rigorous performance evaluation

2. SYSTEM ANALYSIS

2. SYSTEM ANALYSIS

SYSTEM ANALYSIS

System Analysis is the important phase in the system development process. The System is studied to the minute details and analyzed. The system analyst plays an important role of an interrogator and dwells deep into the working of the present system. In analysis, a detailed study of these operations performed by the system and their relationships within and outside the system is done. A key question considered here is, “what must be done to solve the problem?” The system is viewed as a whole and the inputs to the system are identified. Once analysis is completed the analyst has a firm understanding of what is to be done.

2.1 PROBLEM DEFINITION

The challenge is the widespread presence of fake profiles on Facebook, necessitating the development of an artificial intelligence system using machine learning, particularly artificial neural networks, to accurately differentiate between genuine and fake profiles, thus bolstering user security and trust in online interactions.

2.2 EXISTING SYSTEM

Malicious users create fake profiles to phish login information from unsuspecting users. A fake profile will send friend requests to many users with public profiles. These counterfeit profiles bait unsuspecting users with pictures of people that are considered attractive. Once the user accepts the request, the owner of the phony profile will spam friend requests to anyone this user is a friend. The fake profile's contents typically have links that lead to an external website where the damage happens. An unaware curious user clicking the bad link will damage their computer. The cost can be as simple as catching a virus to as bad as installing a rootkit turning the computer into a zombie. While Facebook has a rigorous screening to keep these fake accounts out, it only takes one fake profile to damage the computers of many.

2.2.1 DISADVANTAGES OF EXISTING SYSTEM

- False Positives: Legitimate user accounts can be mistakenly flagged as fake.
- Privacy Concerns: Users may have concerns about their online activities being monitored.
- Adaptability of Fake Profiles: Fake profiles evolve to evade detection methods.
- Limited Transparency: Detection criteria and algorithms are often undisclosed.
- Resource Intensive: Requires significant computational resources and manpower.
- Cat-and-Mouse Game: Ongoing battle as fake account creators adapt.
- User Reporting Bias: Reliance on user reports can lead to false flags and biases.

2.3 PROPOSED SYSTEM

In our solution, we use machine learning, namely an artificial neural network to determine what are the chances that a friend request is authentic or not. We utilize Microsoft Excel to store old and new fake data profiles. The algorithm then stores the data in a data frame. This collection of data will be divided into a training set and a testing set. We would need a data set from the social media sites to train our model. For the training set, the features that we use to determine a fake profile are Account age, Gender, User age, Link in the description, Number of messages sent out, Number of friend requests sent out, Entered location, Location by IP, Fake or Not. Each of these parameters is tested and assigned a value. For example, for the gender parameter if the profile can be determined to be a female or male a value of (1) is assigned to the training set for Gender. The same process is applied to other parameters.

2.3.1 ADVANTAGES OF THE PROPOSED SYSTEM

- Vote Trust uses a voting-based system that pulls user activities to find fake profiles using trust-based vote assignment and global votes total. It is considered as the first line of defence due to limitations which include real accounts that were already compromised being sold.
- Provides a robust tool to identify and mitigate the risks associated with fake profiles, bolstering user security on social media platforms.
- Demonstrates the potential of artificial neural networks and machine learning techniques to address real-world societal challenges, contributing to the advancement of AI.
- Utilizes sophisticated algorithms and feature engineering to accurately identify fake profiles, reducing the burden on platform moderators and improving detection efficiency.

2.4 FEASIBILITY STUDY:

The feasibility of the project is analyzed in this phase and business proposal is put forth with a very general plan for the project and some cost estimates. During system analysis, the feasibility study of the proposed system is to be carried out. This is to ensure that the proposed system is not a burden to the company. For feasibility analysis, some understanding of the major requirements for the system is essential. Three key considerations involved in the feasibility analysis are:

- Economic Feasibility
- Technical Feasibility
- Social Feasibility

2.4.1 ECONOMICAL FEASIBILITY:

This study is carried out to check the economic impact that the system will have on the organization. The amount of fund that the company can pour into the research and development of the system is limited. The expenditures must be justified. Thus the developed system as well within the budget and this was achieved because most of the technologies used are freely available. Only the customized products had to be purchased.

2.4.2 TECHNICAL FEASIBILITY:

This study is carried out to check the technical feasibility, that is, the technical requirements of the system. Any system developed must not have a high demand on the available technical resources. This will lead to high demands on the available technical resources. This will lead to high demands being placed on the client. The developed system must have a modest requirement, as only minimal or null changes are required for implementing this system.

2.4.3 SOCIAL FEASIBILITY:

The aspect of study is to check the level of acceptance of the system by the user. This includes the process of training the user to use the system efficiently. The user must not feel threatened by the system, instead must accept it as a necessity. The level of acceptance by the users solely depends on the methods that are employed to educate the user about the system and to make him familiar with it. His level of confidence must be raised so that he is also able to make some constructive criticism, which is welcomed, as he is the final user of the system.

2.5 HARDWARE & SOFTWARE REQUIREMENTS

2.5.1 HARDWARE REQUIREMENTS:

Hardware interfaces specify the logical characteristics of each interface between the software product and the hardware components of the system. The following are some hardware requirements:

- System : Intel I3 or Above.
- Hard Disk : 100 GB.
- Monitor : 15 inch VGA Color.
- Ram : 4GB or Above.
- CPU :1GHZ.

2.5.2 SOFTWARE REQUIREMENTS:

Software Requirements specifies the logical characteristics of each interface and software components of the system. The following are some software requirements:

- Operating System : Windows XP or above.
- Platform : Python Technology
- Coding Language : Python 3.6
- Tool : Spyder, PyCharm
- Front End : Anaconda
- Back End : Python Anaconda Script

3. ARCHITECTURE

3.1 PROJECT ARCHITECTURE

This project architecture shows the procedure followed for classification, starting from input to final prediction.

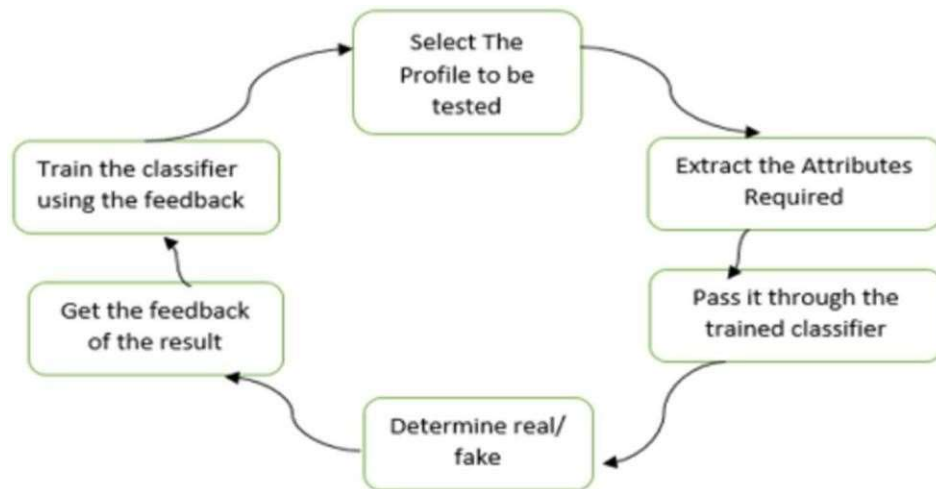


Figure 3.1: Project Architecture for
Use of Artificial Neural Networks to Identify Fake Profiles

3.2 DESCRIPTION

The use of machine learning algorithms such as Random Forest algorithm. Random forest is a supervised learning algorithm that is used for both classifications as well as regression. But however, it is mainly used for classification problems. As we know that a forest is made up of trees and more trees mean more robust forests. Similarly, the random forest algorithm creates decision trees on data samples and then gets the prediction from each of them and finally selects the best solution by means of voting. It is an ensemble method that is better than a single decision tree because it reduces the over-fitting by averaging the result. Each profile (or account) in a social network contains lots of information such as gender, no. of friends, no. of comments, education, work, etc. Some of this information is private and some are public. Since private information is not accessible so, we have used only the information that is public to determine the fake profiles in the social network.

3.3 USE CASE DIAGRAM

In the use case diagram, we have basically one actor who is the user in the trained model.

A use case diagram is a graphical depiction of a user's possible interactions with a system. A use case diagram shows various use cases and different types of users the system has. The use cases are represented by either circles or ellipses. The actors are often shown as stick figures.

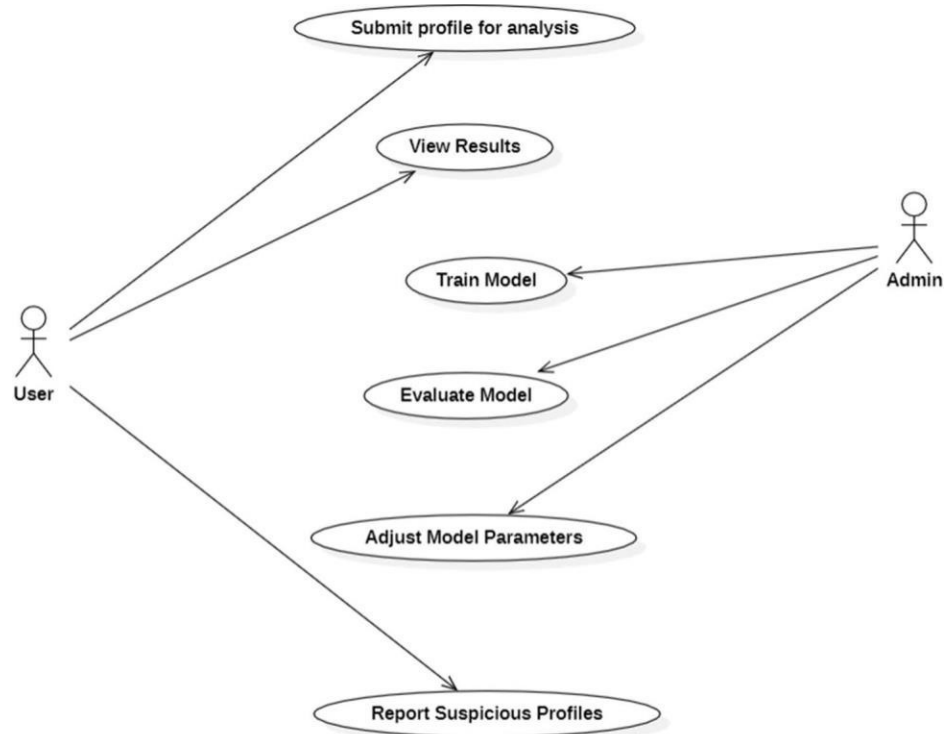


Figure 3.2: Use Case Diagram for Use of Artificial Neural Networks to identify Fake Profiles

3.4 CLASS DIAGRAM

Class diagram is a type of static structure diagram that describes the structure of a system by showing the system's classes, their attributes, operations (or methods), and the relationships among objects.

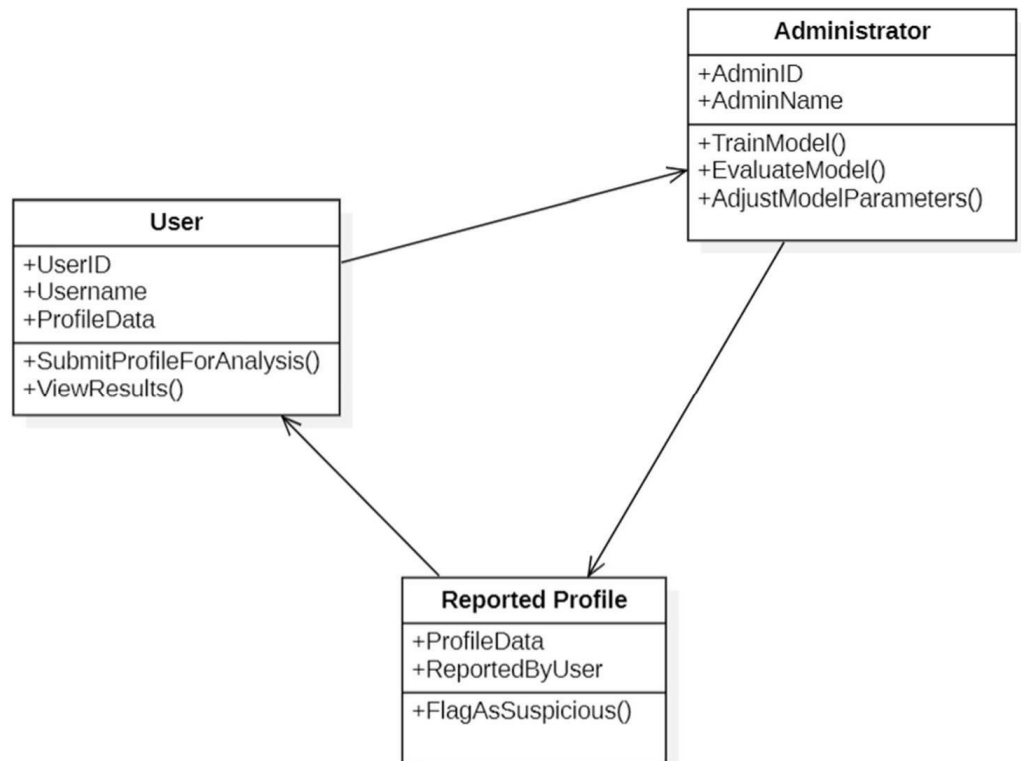


Figure 3.3: Class Diagram for Use of Artificial Neural Networks to identify Fake Profiles

3.4 SEQUENCE DIAGRAM

A sequence diagram shows object interactions arranged in time sequence. It depicts the objects involved in the scenario and the sequence of messages exchanged between the objects needed to carry out the functionality of the scenario. Sequence diagrams are typically associated with use case realizations in the logical view of the system under development.

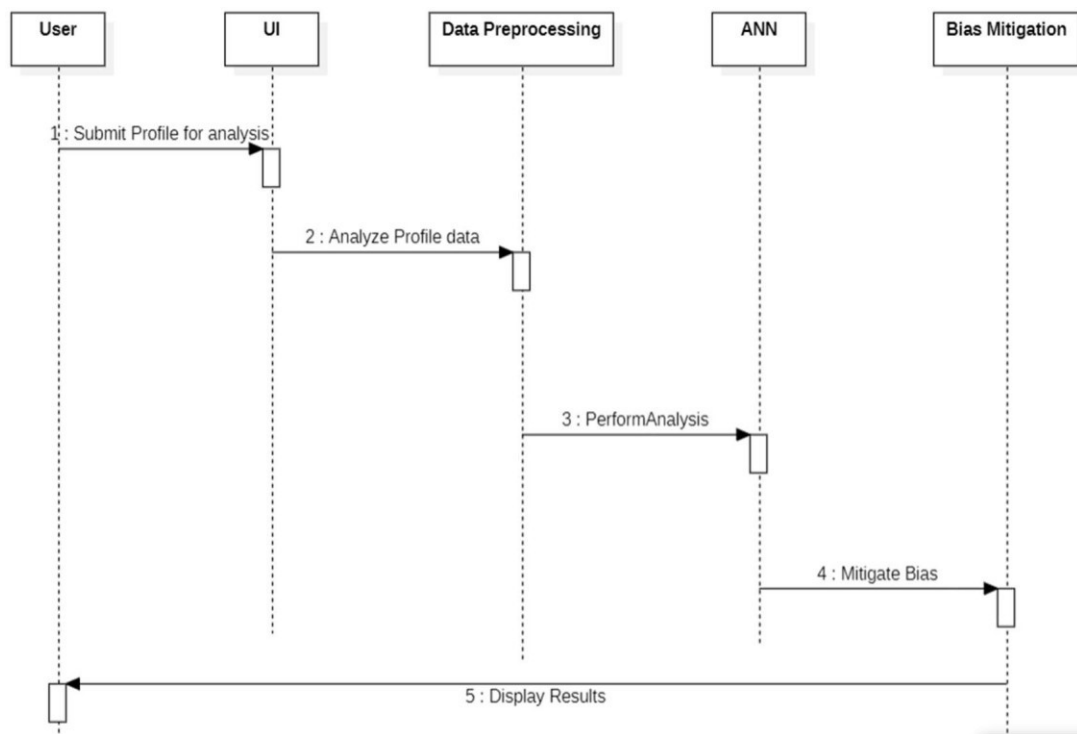


Figure 3.4: Sequence Diagram for
Use of Artificial Neural Networks to identify Fake profiles

3.5. ACTIVITY DIAGRAM

Activity diagrams are graphical representations of workflows of stepwise activities and actions with support for choice, iteration and concurrency. They can also include elements showing the flow of data between activities through one or more data stores.

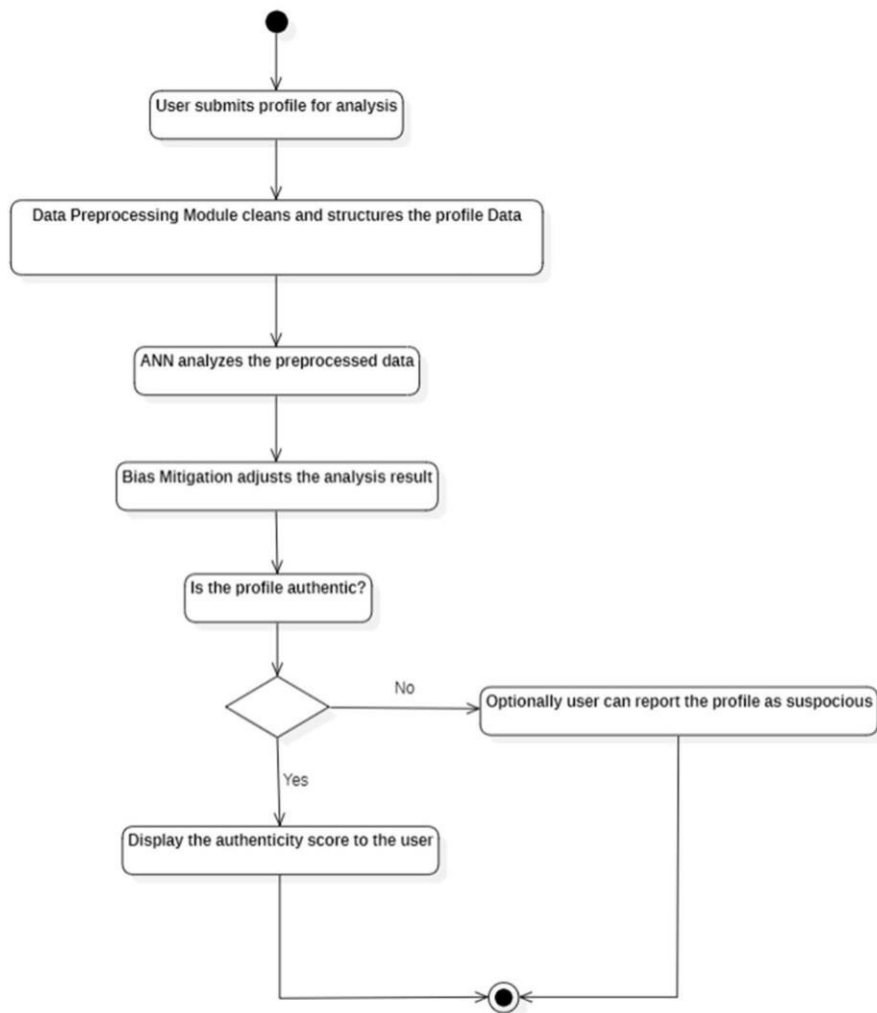


Figure 3.5: Activity Diagram for Use of Artificial Neural Networks to identify Fake profiles

4. IMPLEMENTATION

4.1 SAMPLE CODE

```
from tkinter import messagebox
from tkinter import *
from tkinter import simpledialog
import tkinter
import matplotlib.pyplot as plt
import numpy as np
from tkinter import ttk
from tkinter import filedialog
import pandas as pd
from sklearn.model_selection import train_test_split
from keras.models import Sequential
from keras.layers.core import Dense,Activation,Dropout
from keras.callbacks import EarlyStopping
from sklearn.preprocessing import OneHotEncoder
from keras.optimizers import Adam
from keras.utils.np_utils import to_categorical

main = Tk()
main.title("Fake Account Detection Using Machine Learning and Data Science")
main.geometry("1300x1200")
main.config(bg="lightgreen")
global filename
global X, Y
global X_train, X_test, y_train, y_test
global accuracy
global dataset
global model

def loadProfileDataset():
    global filename
    global dataset
    outputarea.delete('1.0', END)
    filename = filedialog.askopenfilename(initialdir="Dataset")
    outputarea.insert(END,filename+" loaded\n\n")
    dataset = pd.read_csv(filename)
    outputarea.insert(END,str(dataset.head()))
```



```
def preprocessDataset():
    global X, Y global dataset
    global X_train, X_test, y_train, y_test
    outputarea.delete('1.0', END)
    X = dataset.values[:, 0:8]
    Y = dataset.values[:, 8]
    indices = np.arange(X.shape[0])
    np.random.shuffle(indices)
    X = X[indices]
    Y = Y[indices]
    Y = to_categorical(Y)
    X_train, X_test, y_train, y_test = train_test_split(X, Y, test_size=0.2)
    outputarea.insert(END, "\n\nDataset contains total Accounts : "+str(len(X))+"\n")
    outputarea.insert(END, "Total profiles used to train ANN algorithm : "+str(len(X_train))+"\n")
    outputarea.insert(END, "Total profiles used to test ANN algorithm : "+str(len(X_test))+"\n")

def executeANN():
    global model
    outputarea.delete('1.0', END)
    global X_train, X_test, y_train, y_test
    global accuracy
    model = Sequential()
    model.add(Dense(200, input_shape=(8,), activation='relu', name='fc1'))
    model.add(Dense(200, activation='relu', name='fc2'))
    model.add(Dense(2, activation='softmax', name='output'))
    optimizer = Adam(lr=0.001)
    model.compile(optimizer, loss='categorical_crossentropy', metrics=['accuracy'])
    print('ANN Neural Network Model Summary: ')
    print(model.summary())
    hist = model.fit(X_train, y_train, verbose=2, batch_size=5, epochs=200)
    results = model.evaluate(X_test, y_test)
    ann_acc = results[1] * 100
    print(ann_acc)
    accuracy = hist.history
    acc = accuracy['accuracy']
    acc = acc[199] * 100
    outputarea.insert(END, "ANN model generated and its prediction accuracy is : "+str(acc)+"\n")
```

```
def graph():
    global accuracy
    acc = accuracy['accuracy']
    loss = accuracy['loss']
    plt.figure(figsize=(10,6))
    plt.grid(True)
    plt.xlabel('Iterations')
    plt.ylabel('Accuracy/Loss')
    plt.plot(acc, 'ro-', color = 'green')
    plt.plot(loss, 'ro-', color = 'blue')
    plt.legend(['Accuracy', 'Loss'], loc='upper left')
    #plt.xticks(wordloss.index)
    plt.title('ANN Iteration Wise Accuracy & Loss Graph')
    plt.show()

def predictProfile():
    outputarea.delete('1.0', END)
    global model
    filename = filedialog.askopenfilename(initialdir="Dataset")
    test = pd.read_csv(filename)
    test = test.values[:, 0:8]
    predict = model.predict_classes(test)
    print(predict)
    for i in range(len(test)):
        msg = "
        if str(predict[i]) == '0':
            msg = "Given Account Details Predicted As Genuine"
        if str(predict[i]) == '1':
            msg = "Given Account Details Predicted As Fake"
        outputarea.insert(END,str(test[i])+" "+msg+"\n\n")

def close():
    main.destroy()
    font = ('times', 15, 'bold')
    title = Label(main, text='Fake Account Detection Using Machine Learning and Data Science')
    #title.config(bg='powder blue', fg='olive drab')
    title.config(font=font)
    title.config(height=3, width=120)
    title.place(x=0,y=5)
    font1 = ('times', 13, 'bold')
    ff = ('times', 12, 'bold')
```

```
uploadButton = Button(main, text="Upload Social Network Profiles Dataset",
command=loadProfileDataset)
uploadButton.place(x=20,y=100)
uploadButton.config(font=ff)

processButton = Button(main, text="Preprocess Dataset", command=preprocessDataset)
processButton.place(x=20,y=150)
processButton.config(font=ff)

annButton = Button(main, text="Run ANN Algorithm", command=executeANN)
annButton.place(x=20,y=200)
annButton.config(font=ff)

graphButton = Button(main, text="ANN Accuracy & Loss Graph", command=graph)
graphButton.place(x=20,y=250)
graphButton.config(font=ff)

predictButton = Button(main, text="Predict Fake/Genuine Profile using ANN",
command=predictProfile)
predictButton.place(x=20,y=300)
predictButton.config(font=ff)

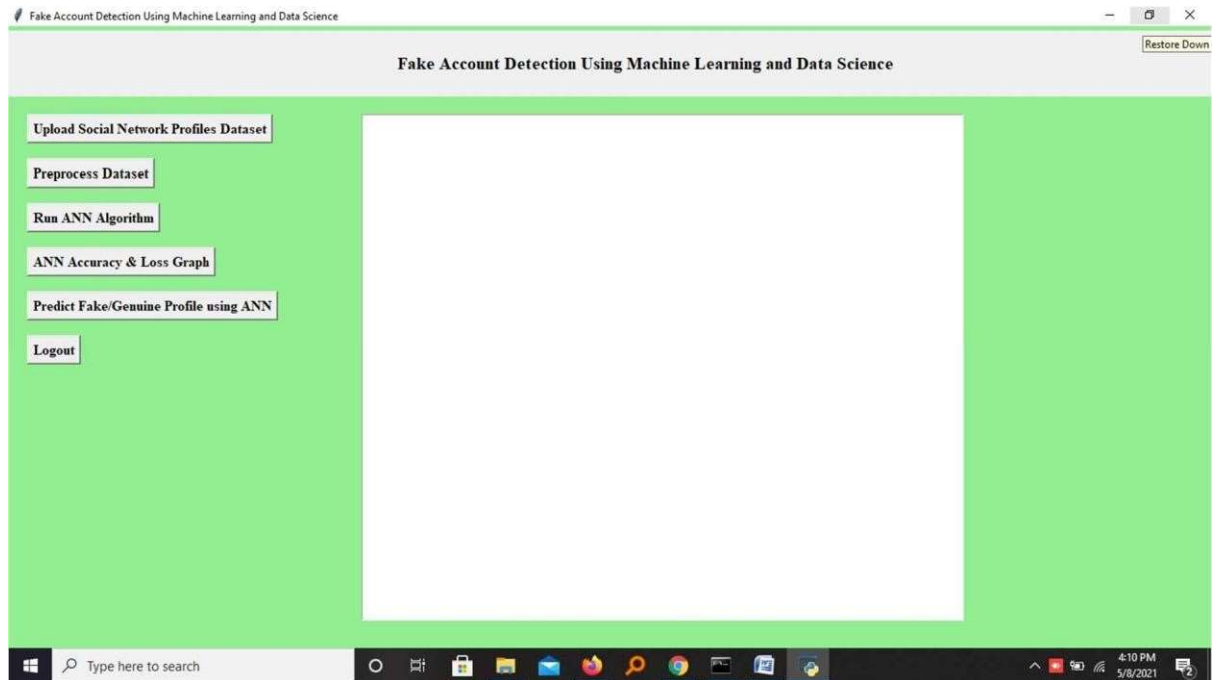
exitButton = Button(main, text="Logout", command=close)
exitButton.place(x=20,y=350)
exitButton.config(font=ff)

font1 = ('times', 12, 'bold')
outputarea = Text(main,height=30,width=85)
scroll = Scrollbar(outputarea)
outputarea.configure(yscrollcommand=scroll.set)
outputarea.place(x=400,y=100)
outputarea.config(font=font1)

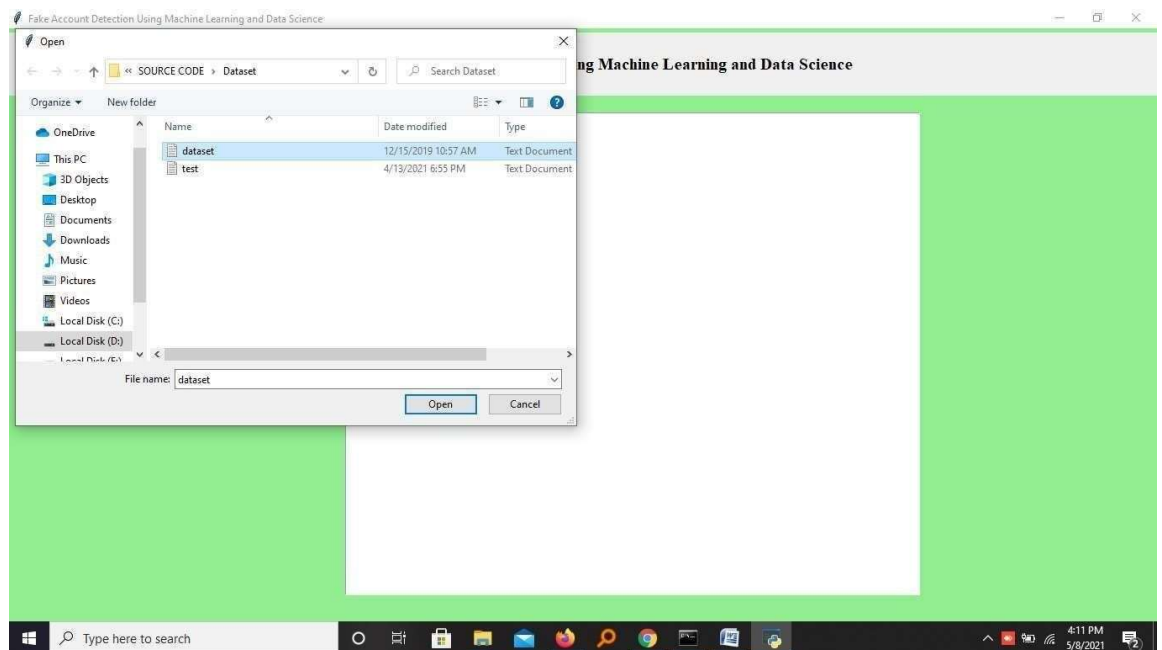
main.config()
main.mainloop()
```

5. SCREENSHOTS

Use Of Artificial Neural Networks To Identify Fake Profiles

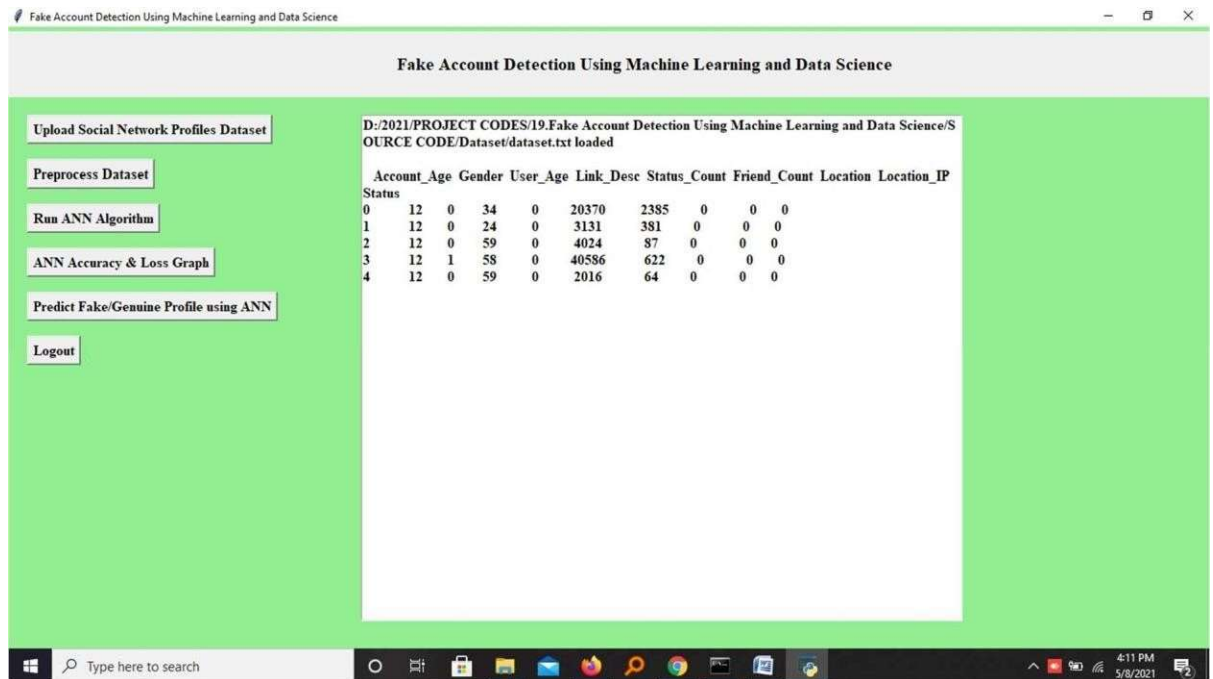


Screenshot 5.1: Upload Social Network Profiles Dataset

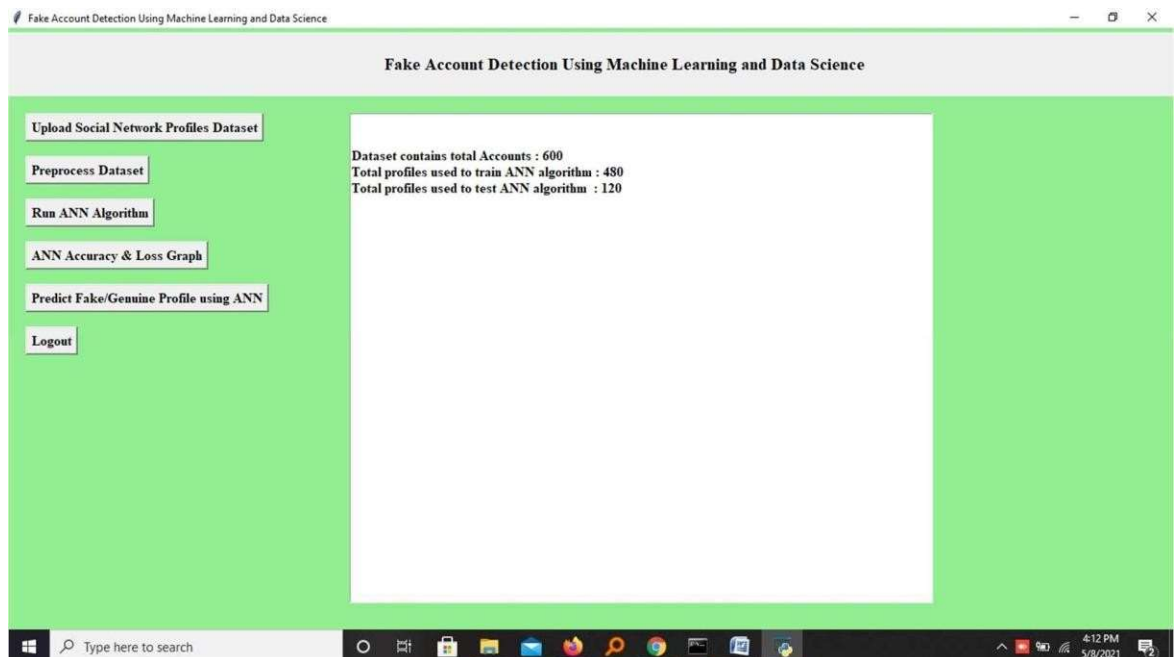


Screenshot 5.2: Select Dataset

Use Of Artificial Neural Networks To Identify Fake Profiles

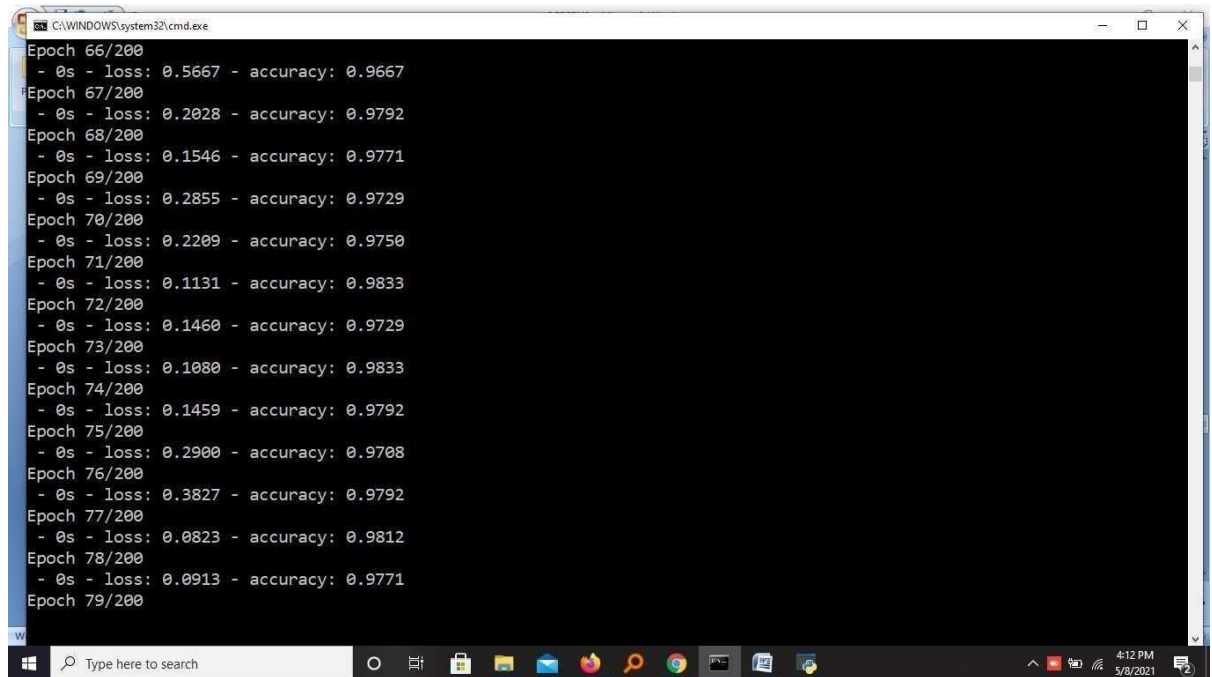


Screenshot 5.3: Preprocess Dataset



Screenshot 5.4: Run ANN Algorithm

Use Of Artificial Neural Networks To Identify Fake Profiles



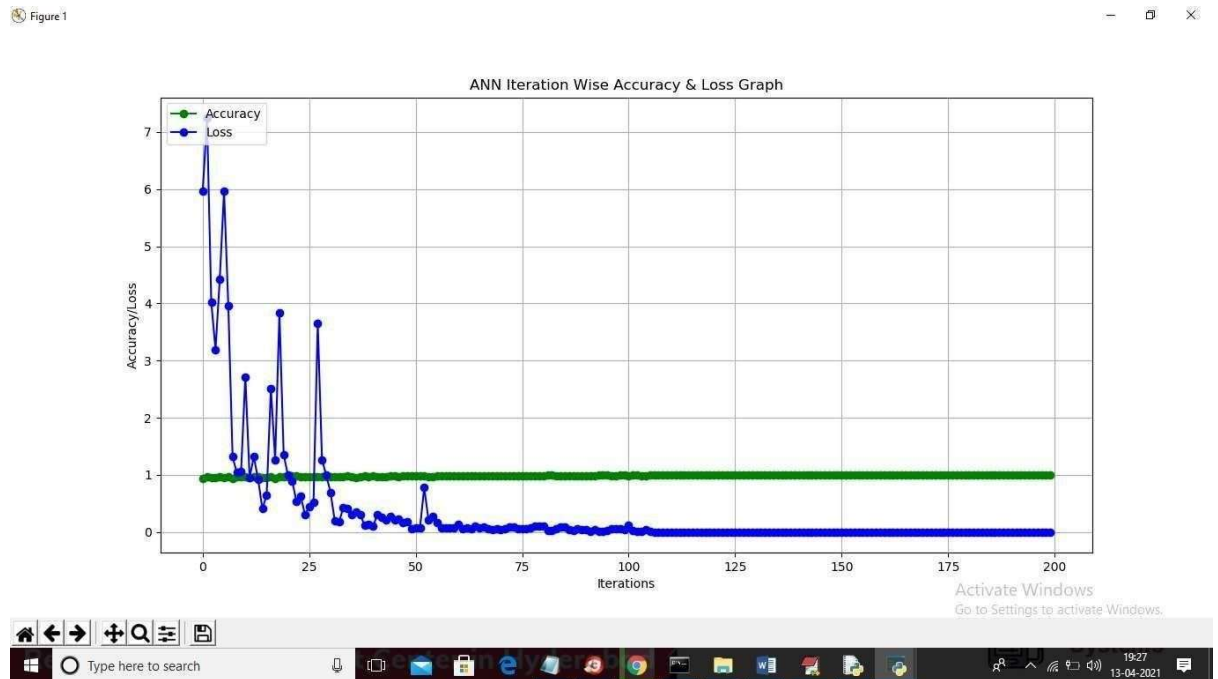
```
C:\WINDOWS\system32\cmd.exe
Epoch 66/200
- 0s - loss: 0.5667 - accuracy: 0.9667
Epoch 67/200
- 0s - loss: 0.2028 - accuracy: 0.9792
Epoch 68/200
- 0s - loss: 0.1546 - accuracy: 0.9771
Epoch 69/200
- 0s - loss: 0.2855 - accuracy: 0.9729
Epoch 70/200
- 0s - loss: 0.2209 - accuracy: 0.9750
Epoch 71/200
- 0s - loss: 0.1131 - accuracy: 0.9833
Epoch 72/200
- 0s - loss: 0.1460 - accuracy: 0.9729
Epoch 73/200
- 0s - loss: 0.1080 - accuracy: 0.9833
Epoch 74/200
- 0s - loss: 0.1459 - accuracy: 0.9792
Epoch 75/200
- 0s - loss: 0.2900 - accuracy: 0.9708
Epoch 76/200
- 0s - loss: 0.3827 - accuracy: 0.9792
Epoch 77/200
- 0s - loss: 0.0823 - accuracy: 0.9812
Epoch 78/200
- 0s - loss: 0.0913 - accuracy: 0.9771
Epoch 79/200
```

Screenshot 5.5: Observe Accuracy

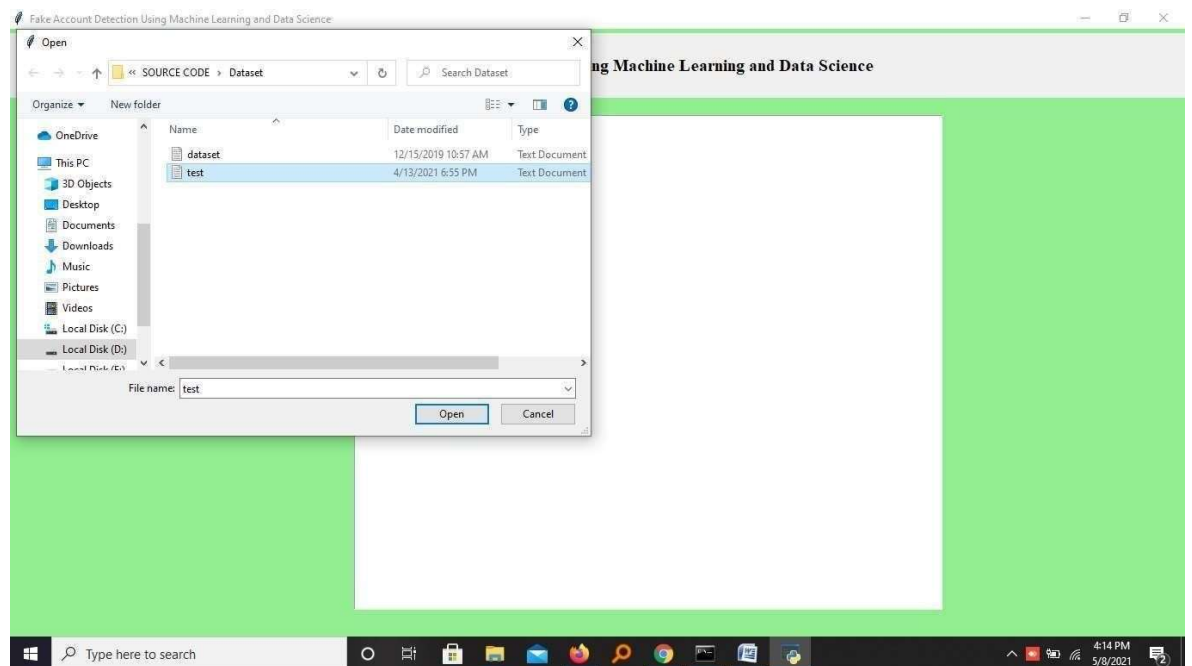


Screenshot 5.6: Generate ANN Accuracy and Loss graph

Use Of Artificial Neural Networks To Identify Fake Profiles

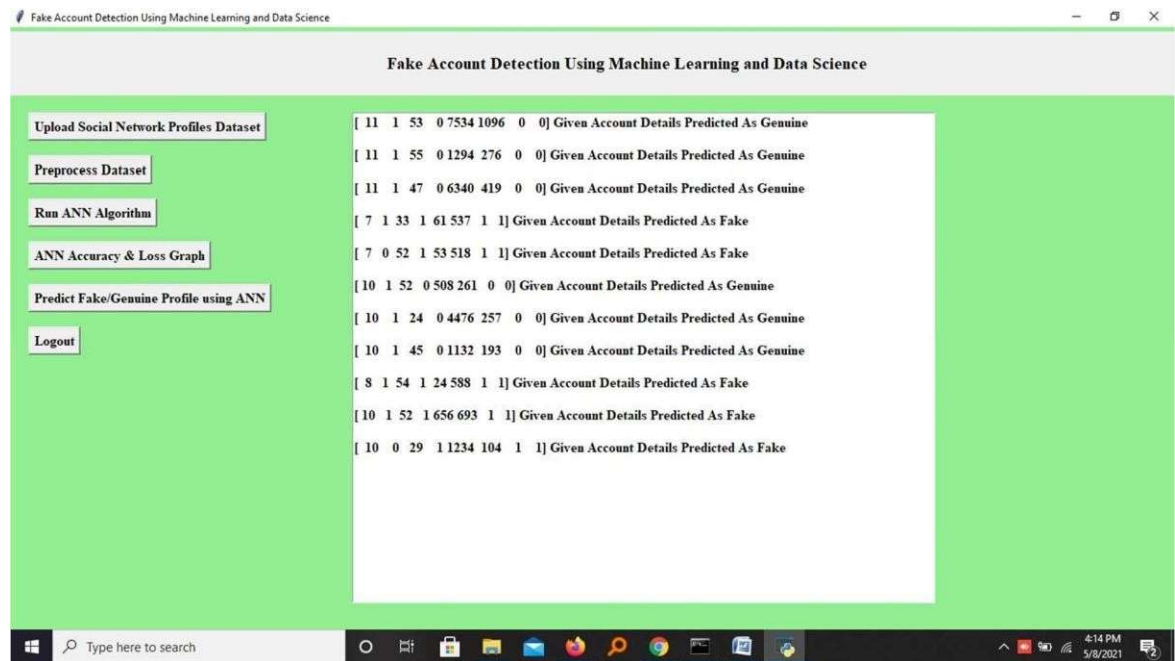


Screenshot 5.7: ANN Accuracy and Loss Graph



Screenshot 5.8: Load Test Data

Use Of Artificial Neural Networks To Identify Fake Profiles



Screenshot 5.9: Fake/Genuine Profile prediction

6. TESTING

6. TESTING

6.1 INTRODUCTION TO TESTING

The purpose of testing is to discover errors. Testing is the process of trying to discover every conceivable fault or weakness in a work product. It provides a way to check the functionality of components, subassemblies, assemblies and/or a finished product. It is the process of exercising software with the intent of ensuring that the Software system meets its requirements and user expectations and does not fail in an unacceptable manner. There are various types of tests. Each test type addresses a specific testing requirement.

6.2 TYPES OF TESTING

6.2.1 UNIT TESTING

Unit testing involves the design of test cases that validate that the internal program logic is functioning properly, and that program inputs produce valid outputs. All decision branches and internal code flow should be validated. It is the testing of individual software units of the application. It is done after the completion of an individual unit before integration. This is a structural testing that relies on knowledge of its construction and is invasive. Unit tests perform basic tests at component level and test a specific business process, application and/or system configuration. Unit tests ensure that each unique path of a business process performs accurately to the documented specifications and contains clearly defined inputs and expected results.

6.2.2 INTEGRATION TESTING

Integration tests are designed to test integrated software components to determine if they actually run as one program. Integration tests demonstrate that although the components were individually satisfactory, as shown by successfully unit testing, the combination of components is correct and consistent. Integration testing is

specifically aimed at exposing the problems that arise from the combination of components.

6.2.3 FUNCTIONAL TESTING

Functional tests provide systematic demonstrations that functions tested are available as specified by the business and technical requirements, system documentation, and user manuals.

Functional testing is centered on the following items:

Valid Input: identified classes of valid input must be accepted.

Invalid Input: identified classes of invalid input must Input be rejected.

Functions: identified functions must be exercised.

Output: identified classes of application outputs must be exercised.

Systems/Procedures: interfacing systems or procedures must be invoked. Organization and preparation of functional tests is focused on requirements, key functions, or special test cases. In addition, systematic coverage pertaining to identify Business process flows; data fields, predefined processes, and successive processes must be considered for testing. Before functional testing is complete, additional tests are identified and the effective value of current tests is determined.

6.3 TEST CASES

6.3.1 CLASSIFICATION

Test Case ID	Test Case Description	Test Case Type	Priority	Preconditions	Steps to Execute	Expected Results
TC001	User submits an authentic profile.	Positive	High	User logged in	1. Submit profile	Authenticity score displayed is high.
TC002	User submits a suspicious profile.	Positive	High	User logged in	1. Submit profile	Authenticity score displayed is low.
TC003	Administrator trains the ANN model.	Positive	High	Administrator logged in	1. Train model	Model training completed successfully.
TC004	Administrator evaluates the model.	Positive	High	Administrator logged in	1. Evaluate model	Evaluation metrics are generated.
TC005	Administrator adjusts model parameters.	Positive	High	Administrator logged in	1. Adjust parameters	Model parameters updated successfully.
TC006	User reports a profile as suspicious.	Positive	Medium	User logged in	1. Report profile	Profile marked as suspicious for review.
TC007	User submits a profile with invalid data.	Negative	Medium	User logged in	1. Submit profile	Error message displayed.
TC008	User views authenticity results.	Positive	Medium	User logged in	1. View results	Authenticity score displayed.
TC009	User interacts with an optional UI.	Positive	Medium	User logged in	1. Use UI features	UI functions as expected.
TC010	Administrator evaluates model fairness.	Positive	Low	Administrator logged in	1. Evaluate fairness	Bias mitigation techniques applied successfully.

7. CONCLUSION

7. CONCLUSION & FUTURE SCOPE

7.1 PROJECT CONCLUSION

We use machine learning, namely an artificial neural network to determine what are the chances that a friend request is authentic or not. Each equation at each neuron (node) is put through a Sigmoid function. We use a training data set by Facebook or other social networks. This would allow the presented deep learning algorithm to learn the patterns of bot behavior by back propagation, minimizing the final cost function and adjusting each neuron's weight and bias.

7.2 FUTURE SCOPE

Multilingual and Cross-Platform Support: Extend the system's capabilities to detect fake profiles in multiple languages and across various social media platforms, making it more versatile and widely applicable.

Continuous Learning and Model Updates: Implement mechanisms for continuous learning and model updates. Incorporate feedback from user reports and evolving tactics of fake profile creators to improve detection accuracy.

Real-Time Profile Analysis: Explore the possibility of real-time profile analysis, enabling users to assess the authenticity of profiles as they encounter them in their online interactions.

Advanced Biometric Authentication: Enhance security by incorporating advanced biometric authentication methods, such as facial recognition, voice recognition, or fingerprint scanning, into the profile authenticity detection process.

Integration with Social Media APIs: Develop integrations with social media platform APIs to offer seamless profile analysis and reporting directly within social media apps, enhancing user convenience.

AI-Based Content Moderation: Expand the system's capabilities to not only detect fake profiles but also identify and moderate harmful content, such as hate speech, fake news, and inappropriate images or posts.

User and Content Reputation Scoring: Introduce a reputation scoring system for both users and the content they share, helping users identify trustworthy connections and reliable information sources.

8. BIBLIOGRAPHY

8.BIBLIOGRAPHY

8.1 REFERENCES

- [1] Kumud Patel; Sudhanshu Agrahari; Saijshree Srivastava “Survey on Fake Profile Detection on Social Sites by Using Machine Learning Algorithm 2020 8th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions) (ICRITO) Year: 2020 | ConferencePaper | Publisher: IEEE
- [2] “Keynote Speech 2: Detecting fake news and profiling fake news spreadersand conspiracy propagators” 2021 12th International Conference on Information and Communication Systems (ICICS) Year: 2021 | Conference Paper | Publisher: IEEE
- [3] Samuel Delgado Muñoz; Edward Paul Guillén Pinto “A dataset for the detection of fake profiles on social networking services” 2020 International Conference on Computational Science and Computational Intelligence (CSCI) Year: 2020 | Conference Paper | Publisher: IEEE
- [4] Pradeep Kumar Roy; Shivam Chahar “Fake Profile Detection on Social Networking Websites: A Comprehensive Review” IEEE Transactions on Artificial Intelligence Year: 2020 | Volume: 1, Issue: 3 | Journal Article | Publisher: IEEE
- [5] P Sowmya; Madhumita Chatterjee “Detection of Fake and Clone accountsin Twitter using Classification and Distance Measure Algorithms” 2020 International Conference on Communication and Signal Processing (ICCSP) Year: 2020 | Conference Paper | Publisher: IEEE

8.2 GITHUB LINK

https://github.com/immubashir/use_of_artificial_neural_networks_to_identify_fake_profiles