Final Year BTech Project-II Report

On

Fake Profile Detection on Social Media

For the Degree of

Bachelor of Technology

In

Computer Science and Engineering

Submitted by

2020BTECS00025 Sahil Otari

2020BTECS00101 Kranti Bharti

2020BTECS00029 Akshata Gawande

2020BTECS00044 Sanjana Kondalwade

Under the Guidance of

Prof. M.K. Chavan

Department of Computer Science and Engineering,

Walchand College of Engineering, Sangli.

(*An Autonomous Institute*)

AY 2023-24

Walchand College of Engineering, Sangli.

(*An Autonomous Institute*)

Department of Computer Science and Engineering

CERTIFICATE

This is to certify that the Project Report entitled, **"FAKE PROFILE DETECTION ON SOCIAL MEDIA "** submitted by

2020BTECS00025 Sahil Otari
2020BTECS00101 Kranti Bharti
2020BTECS00029 Akshata Gawande
2020BTECS00044 Sanjana Kondalwade

to **Walchand College of Engineering, Sangli**, India, is a record of bonfire Project work of course **"PROJECT-II (5CS492)"** carried out by him/her under my/our supervision and guidance and is worthy of consideration for the award of the degree of Bachelor of Technology in Computer Science & Engineering during the academic year ***2023-24***.

Prof. M. K. Chavan
Guide

External Examiner

Mrs. Dr. M. A. Shah
Head
Department of Computer Science and Engineering

Declaration

I hereby declare that work presented in this project report titled **"FAKE PROFILE DETECTTION ON SOCIAL MEDIA"** submitted by us in the partial fulfilment of the requirement of the award of the degree of Bachelor of Technology (B.Tech) Submitted in the Department of Computer Science & Engineering, Walchand College of Engineering, Sangli, is an authentic record of my project work carried out under the guidance of Prof. M. K. Chavan.

2020BTECS00025 Sahil Otari

2020BTECS00101 Kranti Bharti

2020BTECS00029 Akshata Gawande

2020BTECS00044 Sanjana Kondalwade

Place : Sangli. Date : 13/05/2024

Acknowledgement

We would like to express our sincere gratitude to all those who have contributed to the completion of this college project report.

First and foremost, we extend our deepest appreciation to our project guide Prof. M. K. Chavan, whose invaluable guidance, support, and encouragement throughout the project were instrumental in its success. He provided us with expert advice, constructive feedback, and insightful suggestion, which greatly enriched our understanding and helped us overcome various challenges.

We are also grateful to our HOD Dr. Medha Shah and faculty members for their valuable inputs, encouragement, and academic support throughout the duration of the project. Their expertise and guidance have been invaluable in shaping our project and enhancing its quality.

We extend our thanks to our classmates and friends for their encouragement, assistance, and constructive feedback during the course of this project. Their support and encouragement have been a constant source of motivation for us

Last but not least, we would like to express our heartfelt gratitude to our families for their unwavering support, understanding, and encouragement throughout this endeavour. Their love, encouragement, and patience have been the driving force behind our success.

Thank you all for your invaluable contributions to this project.

Abstract

The rise of social networking has changed how people connect online, with platforms like Twitter, Facebook, and Instagram allowing users to share information easily. However, many users are unaware of the security risks associated with these platforms, including fraudulent profiles and online identity theft. Fake profiles can spread misinformation or engage in malicious activities, posing a threat to users' privacy and security. In addition to the security threats posed by fake profiles, modern online social networks face various other challenges. These include issues like privacy concerns, cyberbullying, and the spread of misinformation leading to individual harm.

To tackle this issue, various machine learning and deep learning techniques like Random Forest, Support Vector Machine, Gradient Boost Classifier, Logistic Regression, Artificial Neural Network, and others can be utilized to detect fake profiles. Among these models, random forest has better result compared to other machine learning models. But these models were not able to detect the profile accurately as fake or genuine.

This project explores and compares the effectiveness of these techniques in identifying and combating the rapid growth of fake profiles on social networks. Despite the widespread use of social media, many users remain unaware of these risks, highlighting the need for effective solutions. By leveraging machine learning models, this project aims to address these challenges and enhance the safety and security of online social networks. This project evaluates the efficiency of different machine learning techniques and deep learning techniques in addressing these issues and provides an overview into the future of social network security.

CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1: INTRODUCTION

1.1 Introduction:

The new era of networking has started due to easy accessibility of various social networking applications. People are sharing huge amount of data using social network platforms which includes Instagram, Twitter, Facebook and so on. Despite the fact that social networks proves to be very useful to us, but it has some drawbacks also. There are several issues with modern online social networks, such as fraudulent profiles and online impersonation. Social media is facing a major challenge with the rise of fake profiles, which are essentially accounts pretending to be someone they're not. These profiles engage in dishonest activities, causing disruptions and unwanted behaviour on social networks. They spread false information about individuals or even attempt fraudulent actions with harmful intentions. Fortunately, machine learning methods offer a way to detect these fake profiles. This project examines and compares various machine learning approaches aimed at identifying fake profiles to mitigate these issues.

Data of the people who have their active accounts has been considered for the dataset. That dataset has been going for further processing. The same has been explained in the flow diagram. In this case, the data we've gathered undergoes preparation before it enters the classification phase. This step, known as data pre-processing, is crucial for ensuring the accuracy of the classifier's results. Various categorization models are applied to sort the collected data into categories of fake or real profiles. Initially, the model is trained using available data as part of the classification process. Then, it's used to predict whether a user account is fake or real, considering data that wasn't part of the initial dataset. We can evaluate the performance of the model using parameters such as accuracy, precision, recall, and F1-score.

Machine Learning and Deep Learning models enable systems to learn from data, recognize patterns, and draw conclusions with minimal human involvement. Many studies utilize machine learning methods to detect fake accounts, as indicated in the literature. Support Vector Machine, Naive Bayes, Random Forest, Gradient Boost Classifier and Sequential Model are among the commonly used machine learning techniques for identifying fake profile.

In our project report, we explore various machine learning models to detect fake profiles on social media platforms. We have used Support Vector Machine (SVM), Random Forests, K-Nearest Neighbours (KNN), Gradient Boost Classifier, Logistic Regression, Multi Layer Perceptron, Naive Bayes models that predicts the probability of a profile being fake or real based on its variety of features.

In addition to our exploration of machine learning models for detecting fake profiles on social media, we also highlight the ethical considerations surrounding such technologies. As these algorithms become more sophisticated, there's a growing concern about and privacy implications. We acknowledge the importance of responsible Artificial Intelligence and Deep Learning development and emphasize the need for transparency and fairness in deploying these models. By addressing these concerns upfront, we have try to contribute in a more comprehensive way of the implications by using machine learning and deep learning in social media security.

1.2 Literature Review

The 'Rajdavindar' et.al, (2022) that is proposed the literature in the title of paper, which is the need for developing a system that is able to identify social sharing profile by the help of Graph and User based features that have been mostly used. And suggest to improve the accuracy by the using of Decision Tree (DT), Naives Baiye's(NB),and SVM Further. [1]

In another study 'Ananya B' et.al, (2021), To Evaluate a model and applied a lot of Machine Learning Methods to detect authenticity on Social Media Platform and to find the highest accuracy in detecting false profiles, So, the prediction made 93% fake and 96% genuine account correctly. [2]

Yuval et al. [3] are inventing a method for protecting the privacy of users on online social networks. In this invention, they took the fake and legitimate profile data of the existing social site and then feature is extracted from the pre-determined set of the target fake and legitimate profile, the extract feature set chosen from the friend's list and followers of the target profiles. If examining the relationship of each private node and between the social accounts. Classifiers are applied to other existing features of fake profile, by using trained and non-structured supervised machine learning.

The ‘T. Om P’ et.al, (2021), that is used the Machine learning algorithms to detect the fake profile accurately and collecting the dataset like Training dataset (70%), Validating dataset (10%) and Testing Dataset (20%) for each features matrix fed into such as [LR, XGB, ADB, GBM] used to detect the fraudulent account and better achieving accuracy up to 95%. [4]

A theoretical methodology for identifying phony accounts is being explained by Suheel Yousuf et al. [5]. The suggested project makes use of machine learning algorithms like ANNs, Decision Trees, Naïve Bayes, and SVM. supervised machine learning algorithm, advanced noise reduction, and data standardization approach data set were used for the detection of fraudulent profiles. The dataset's nonsignificant at-tributes are found and attribute reduction is performed using the artificial bee colony and ant colony optimization algorithm, which are inspired by nature. The algorithm of the individual support vector machine was implemented for both phony and real profiles. To provide a more accurate prediction in this case, the ensemble classifier is employed.

Preethi Harris et al. [6] deals with Fake Instagram Profile Identification and Classification using Machine Learning models. The fake profile detection process first starts with the selection of the profile that needs to be tested. After selection of the profile, the suitable attributes is selected. The attributes extracted is passed to the trained classifier to determine profiles that are fake and genuine. The classifier gets trained regularly as new training data is feed into the classifier. The classifier determines the accuracy of the profile. Since different types of classification algorithms and clustering algorithm are used the respective accuracy and execution time of all the algorithms is compared. Finally, when the user input is given with all the necessary attributes, it detects whether the profile is fake or genuine.

Jyothis Joseph et al. [7] deals with the Fake Profile Detection in Online Social Networks using Machine Learning Models. These Machine Learning Models includes Support Vector Machine (SVM), K-Nearest Neighbour (KNN), Decision Tree, Random Forest, Naïve Bayes.

Dr.M.Kathirava et al. [8] analyze and detect Fake Profile Over Social Media using Machine Learning Techniques. The goal of this work is to present a model that can distinguish between fraudulent and real accounts using machine learning classification techniques such as Neural

Network (NN) and Random Forest Classifier (RFC). A system of automatic detection. It try to provide a framework for automating the identification of fraudulent profiles, making it easier for sites to deal with the overwhelming number of profiles they must deal with on a daily basis without sacrificing the quality of the user experience.

Ms. S. R. Ramya AP/CSE et al. [9] talks about Fake Account Sign-In Detection using Support Vector Machine (SVM). Fake accounts are identified and prevented in the browser. The technique uses two algorithms: Neural Network (NN) and Support Vector Machine (SVM), to classify the fake accounts and block them in the browser.

Kunam Umbrani at el. [10] works with Fake Profile Detection using Machine Learning. The data uploaded on social media contains their personal information, thoughts on certain topics, news, etc. The social media platforms verify the authenticity of the registered user. However, some of the users hide their identities and these people are threats to the security of other users' data. These bot accounts are used to scam, or purposefully cause harm to people. There is a need for detection techniques to find and eradicate these bots as quickly as possible. In this work, they have proposed a Machine learning based model that can identify fake or bot created accounts accurately.

Jai Gangan at el. [11] helps to detect Fake Twitter Accounts using Ensemble Learning Model. The contribution of this research work lies in providing a robust and reliable solution to combat the proliferation of fake accounts on Twitter. By enabling the timely detection and elimination of fraudulent profiles, our approach strengthens the network's security and protects users from malicious entities, including trolls, internet fraudsters, deceptive advertising campaigns, and sexual predators.

1.3 Problem Statement:

To detect the fake profiles on social media using Machine Learning and Deep Learning.

1.4 Significance:

Fake profiles are often used to spread misinformation, manipulate public opinion, or engage in fraudulent activities. These can be used to gather personal information or engage in identity theft. By identifying and removing such profiles, the privacy and security of

genuine users are better protected. Social media platforms thrive on trust. Users expect that the profiles they interact with are genuine. Detecting and removing fake profiles helps maintain trust among users, leading to a safer and more authentic online environment. Hence this project plays a important role in detecting these fake profiles present on social media.

1.5 Research Objectives:

1. To collect and analyze dataset of fake profiles on social media.
2. To study various machine learning and deep learning algorithms.
3. To build a model that detects fake profiles on social media.
4. To leverage machine learning and deep learning to achieve high level accuracy and to identify proper insights and patterns.

CHAPTER 2: PROPOSED METHODOLOGY

2.1. General Methodology

In this study, a comprehensive approach for Fake profile detection on social media is presented as illustrated in Fig. 1, depicting the workflow diagram detailing data collection, preprocessing, model exploration, fine-tuning, and integration with a web interface.

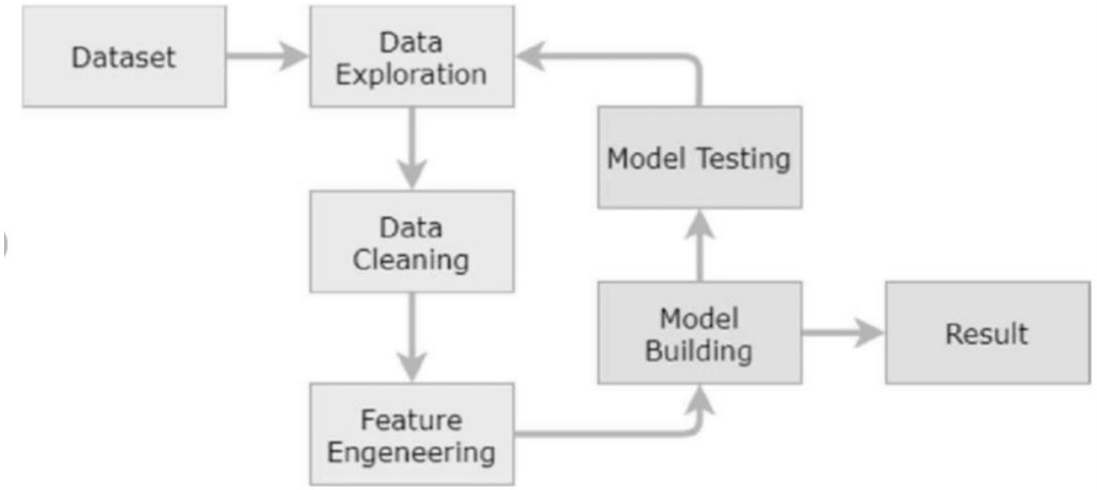

Fig. 1. Workflow Diagram

2.2. Data Collection

The investigation began with a meticulous data collection process focused on uncovering instances of fake profiles across various social media platforms. We meticulously scoured Kaggle's extensive dataset repository to compile a diverse collection of datasets specifically addressing fake profile occurrences. Here the dataset used is Instagram spammer genuine accounts dataset from kaggle. The data set consisted of real profiles and fake ones. This dataset has features like profile length, account type, post, etc. The information is kept in CSV format for machine extraction. Table I below provides an overview of the total number of entries included in the dataset and their respective distributions.

2.3. Data Cleaning

Upon acquiring the datasets, a critical phase of data cleaning is embarked to ensure the integrity and reliability of our dataset. A rigorous cleaning process is implemented to eliminate noise and irrelevant information, thereby enhancing the quality of the dataset for subsequent analysis. This thorough cleaning endeavor was essential in preparing the dataset for further processing and analysis, laying a solid foundation for effective fake profile detection on social media.

2.4. Preprocessing

Following the data cleaning, the preprocessing of textual data to standardize and homogenize it for fake profile detection is examined. The process begins with importing necessary libraries and loading the dataset into training and testing sets. After performing exploratory data analysis (EDA) on both datasets, the data is prepared for feeding into the model by separating features and labels and scaling the data. Then, a simple deep learning or machine learning model is built using TensorFlow, Scikit-learn and Keras. The model is trained using the training data, and its performance is evaluated using validation data. Finally, the model's performance is assessed using metrics such as loss progression during training/validation, classification report, ROC curve and confusion matrix visualization.

2.5. Model Selection

Selecting an appropriate model for fake profile detection posed a significant challenge, necessitating a thorough evaluation of various machine learning and deep learning architectures. A diverse array of models was explored, considering factors such as model performance, computational efficiency, and scalability. The objective was to identify the most effective approach for accurately detecting fake profiles, ensuring optimal model performance. Fig 2. depicts the model selection process employed.

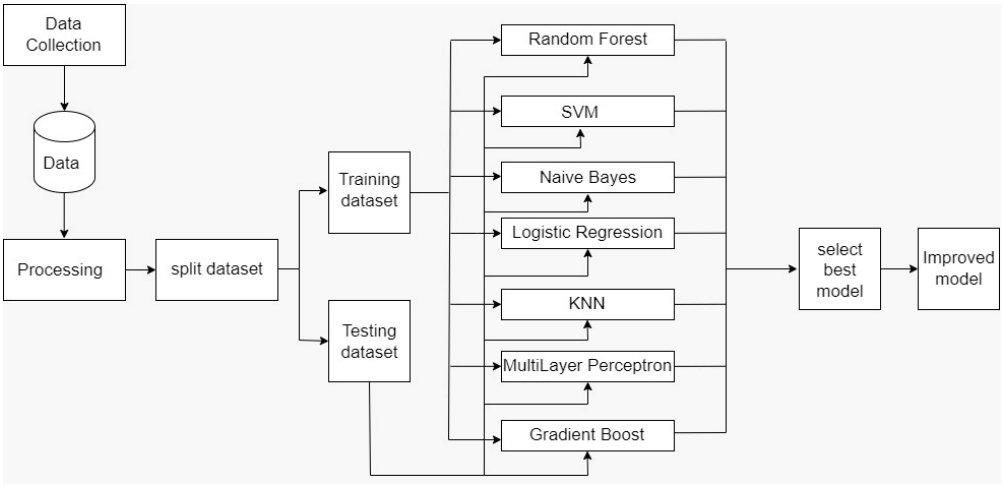

Fig. 2. Model Selection Process

A comprehensive comparative analysis of different models' performance to identify the most effective approach for fake profile classification is conducted. This comparative analysis involved a holistic consideration of each model's classification performance, computational efficiency, and scalability. Insights gleaned from this analysis were instrumental in elucidating

the relative strengths and weaknesses of each model architecture, guiding subsequent decision-making processes regarding model selection, refinement, and optimization. By conducting a meticulous comparison and analysis of different model architectures, the aim was to identify the most effective approach for fake profile detection on social media, ensuring optimal model performance.

A Random Forest

Random Forest is a popular machine learning algorithm that belongs to the supervised learning technique. It can be used for both Classification and Regression problems in ML. It is based on the concept of ensemble learning, which is a process of combining multiple classifiers to solve a complex problem and to improve the performance of the model. We use random forest as it takes less training as compared to other algorithms. It predicts outputs with high accuracy, even for the large dataset it runs effectively. It can also maintain accuracy when a large proportion of data is missing. Random Forest works in two-phase first is to create the random forest by combining N decision tree, and second is to make predictions for each tree created in the first phase. The results from several decision trees applied to various subsets of the input data set are averaged by the Random Forest classifier in order to improve the accuracy of the input data set. The random forest employs predictions from all the trees instead of just one in order to anticipate the result depending on which forecasts received the most votes. Below Fig 3. confusion matrix of random forest and Fig 4. ROC curve of Random Forest is implemented in this project.

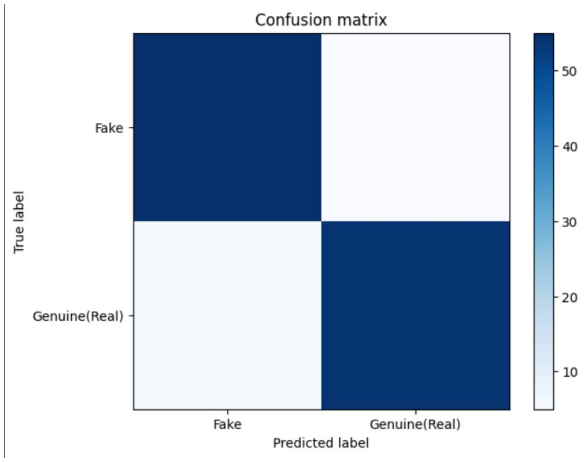

Fig. 3. Confusion Matrix of Random Forest

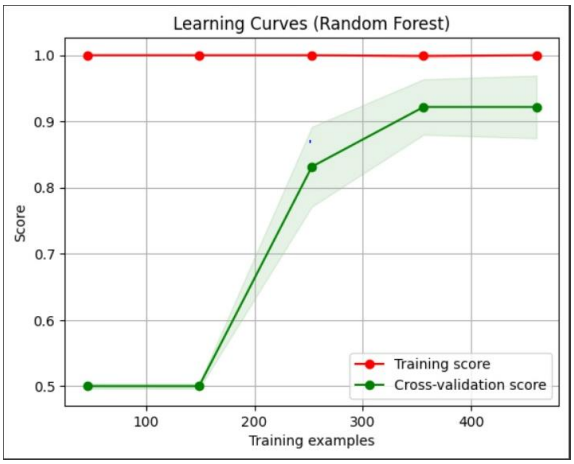

Fig. 4. ROC Curve of Random Forest

B. Support Vector Machine

Support Vector Machine (SVM) is a powerful machine learning algorithm used for linear or nonlinear classification, regression, and even outlier detection tasks. SVMs can be used for a variety of tasks, such as text classification, image classification, spam detection, handwriting identification, gene expression analysis, face detection, and anomaly detection. The main objective of the SVM algorithm is to find the optimal hyperplane in an N-dimensional space that can separate the data points in different classes in the feature space. The hyperplane tries that the margin between the closest points of different classes should be as maximum as possible. The SVM approach aims to identify the optimal line or decision boundary that effectively separates classes within an n-dimensional space, allowing it to categorize new data points in the future. SVM utilizes support vectors, which are extreme data points and vectors to determine the hyperplane that effectively separates classes. The training set of data is processed using a kernel function to convert a non-linear decision surface into a linear equation in a high dimensional space. The kernel used by the SVM for training is linear kernel. Below Fig 5. depicts the confusion matrix and Fig 6. Depicts the ROC curve of support vector machine implemented in this project.

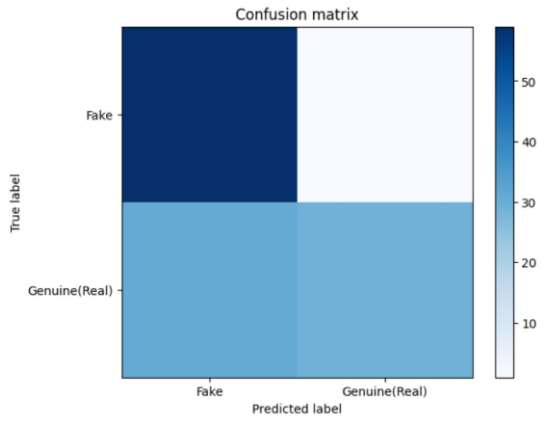

Fig. 5. Confusion Matrix of SVM

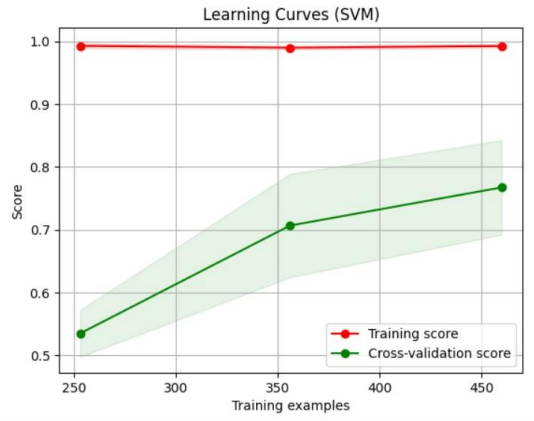

Fig. 6. ROC Curve of SVM

C K-Nearest Neighbour

K-Nearest Neighbour is one of the simplest Machine Learning algorithms based on the Supervised Learning technique. The algorithm assumes the similarity between the new case/data and available cases and puts the new case into the category that is most similar to the available categories. The KNN algorithm stores all the available data and classifies a new data point based on the similarity. This means when new data appears then it can be easily classified

into a well suite category by using the algorithm. K-NN algorithm can be used for Regression as well as for Classification but mostly it is used for Classification problems. A data point is classified using the k nearest neighbour algorithm based on how its neighbours are grouped. The number of nearest data points to include in the majority voting process is indicated by the k value. The value of k is chosen as three. Fig. 7 demonstrates the confusion matrix and Fig 8. demonstrates the ROC curve of K-Nearest neighbour algorithm implemented in this project.

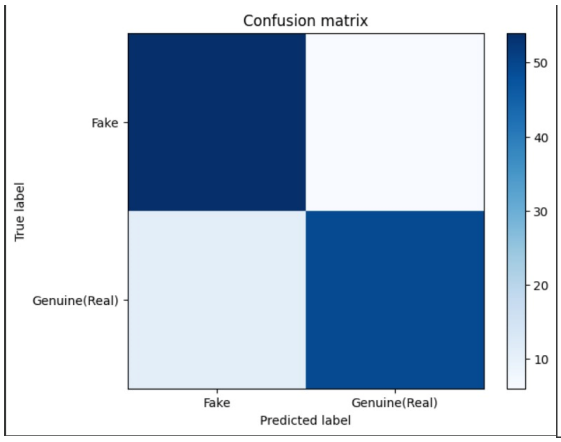

Fig. 7. Confusion Matrix of KNN

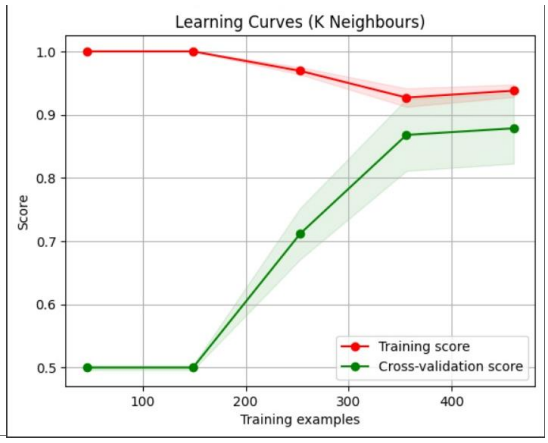

Fig. 8. ROC Curve of KNN

D Multi-layer Perceptron

Multi-layer perception is also known as MLP. It is fully connected dense layers, which transform any input dimension to the desired dimension. A multi-layer perception is a neural network that has multiple layers. To create a neural network we combine neurons together so that the outputs of some neurons are inputs of other neurons. A multi-layer perceptron has one input layer and for each input, there is one neuron(or node), it has one output layer with a single node for each output and it can have any number of hidden layers and each hidden layer can have any number of nodes. When neurons are connected to form a neural network, their outputs become the inputs of other neurons. It is composed of numerous layers that are closely coupled and dense, allowing it to convert any input dimension into any output dimension. The Fig 9. shows the confusion matrix and Fig 10. Shows confusion matrix of Multi-layer perception.

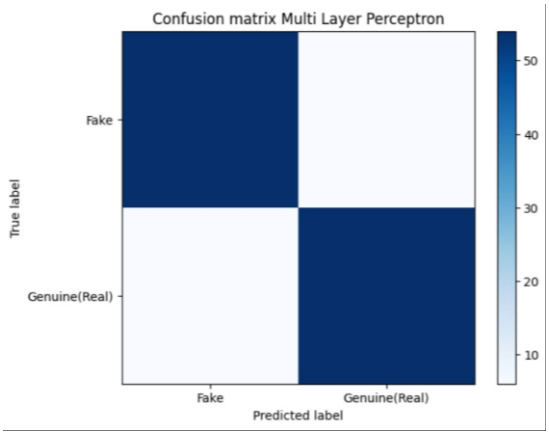

Fig. 9. Confusion Matrix of Multi Layer Perceptron

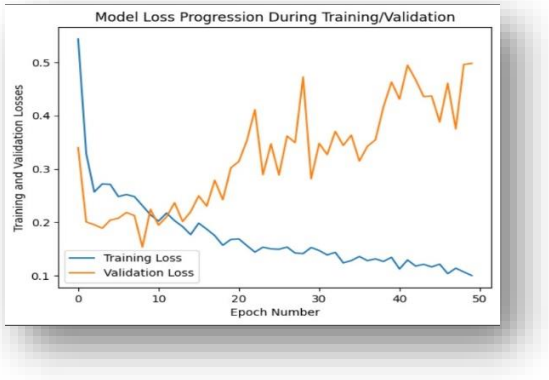

Fig. 10. ROC Curve of Multi Layer Perceptron

E Naive Bayes Classifier

Naïve Bayes algorithm is a supervised learning algorithm, which is based on Bayes theorem and used for solving classification problems. It is mainly used in text classification that includes a high-dimensional training dataset. Naïve Bayes Classifier is one of the simple and most effective Classification algorithms which helps in building the fast machine learning models that can make quick predictions. It is a probabilistic classifier, which means it predicts on the basis of the probability of an object. Continuous values connected to each feature in Gaussian Naive Bayes are distributed in a Gaussian manner. Naive Bayes Classifier algorithm is very easy to implement and computes efficiently, it is effective in cases with a large number of features and it performs well even with limited training data. Here we have dataset of profiles on social media and the target is to find if the profile is fake or real based on the features of the fake profile present it the dataset Below Fig 11. shows confusion matrix and Fig 12. shows ROC curve and of Navie bayes Classifier implemented in this project.

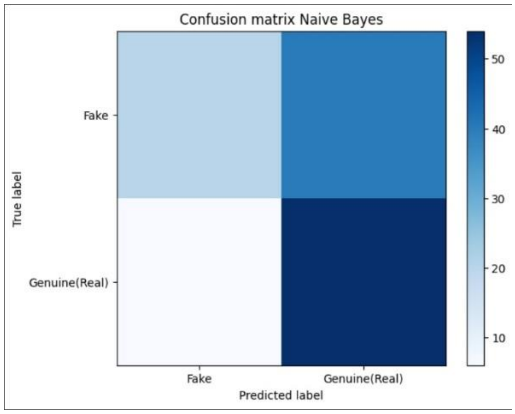

Fig. 11. Confusion Matrix of Naïve Bayes

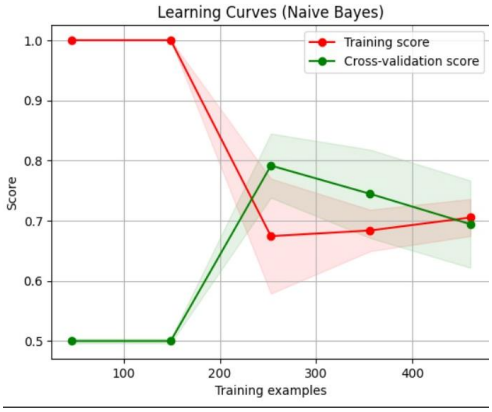

Fig. 12. ROC Curve of Naive Bayes

F Logistic Regression

Logistic regression is a supervised machine learning algorithm used for classification tasks where the goal is to predict the probability that an instance belongs to a given class or not. Logistic regression is a statistical algorithm which analyse the relationship between two data factors. Logistic regression predicts the output of a categorical dependent variable. Therefore, the outcome must be a categorical or discrete value. It can be either Yes or No, 0 or 1, true or False, etc. but instead of giving the exact value as 0 and 1, it gives the probabilistic values which lie between 0 and 1. In Logistic regression, instead of fitting a regression line, we fit an "S" shaped logistic function, which predicts two maximum values (0 or 1). The likelihood that a given instance belongs to a specific class is predicted via logistic regression. The solver used is lbfgs. It has to reject previous gradients and accumulate only new gradients in accordance with the memory restriction because it uses less memory than a regular bfgs. The below Fig.13 depicts the confusion matrix and Fig 14. depicts the ROC curve implementation of logistic regression.

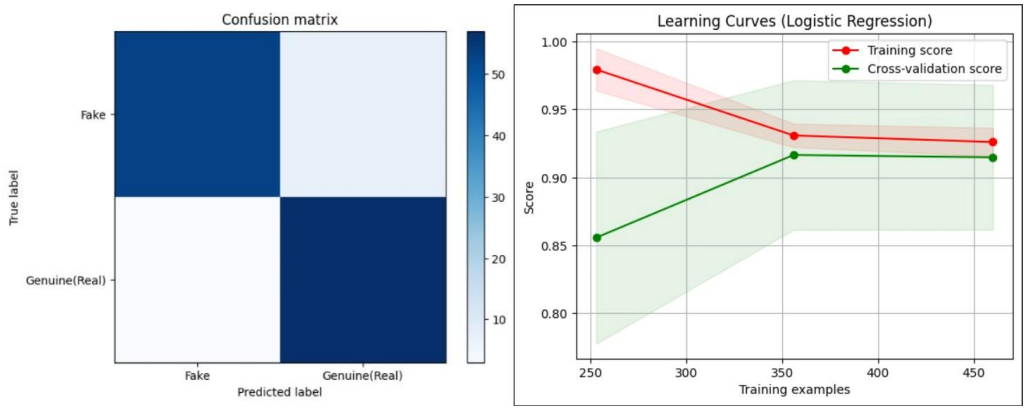

Fig. 13. Confusion Matrix of Logistic Regression Fig. 14. ROC Curve of Logistic Regression

G Gradient Boost Classifier

Gradient Boosting is a powerful boosting algorithm that combines several weak learners into strong learners, in which each new model is trained to minimize the loss function such as mean squared error or cross-entropy of the previous model using gradient descent. In each iteration, the algorithm computes the gradient of the loss function with respect to the predictions of the current ensemble and then trains a new weak model to minimize this gradient. The predictions of the new model are then added to the ensemble, and the process is repeated until a stopping criterion is met. Gradient Boosting updates the weights by computing the negative gradient of the loss function with respect to the predicted output. Gradient Boosting is generally more

robust, as it updates the weights based on the gradients, which are less sensitive to outliers. Below are the results of the algorithm Gradient Boost classifier implemented in this project. Fig 15. demonstrates confusion matrix of Gradient Boost and Fig 16. demonstrates ROC curve of Gradient Boost.

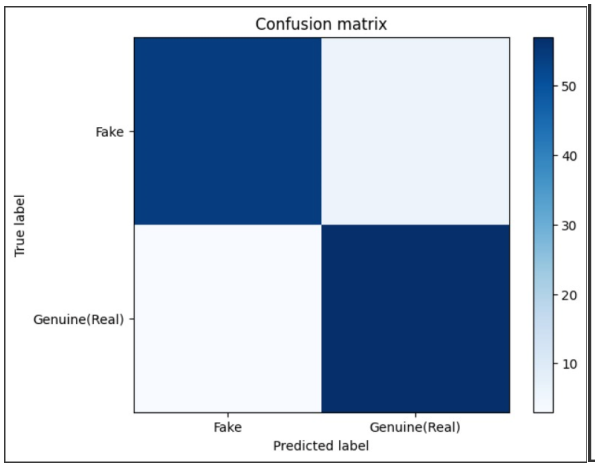

Fig. 15. Confusion Matrix of Gradient Boost

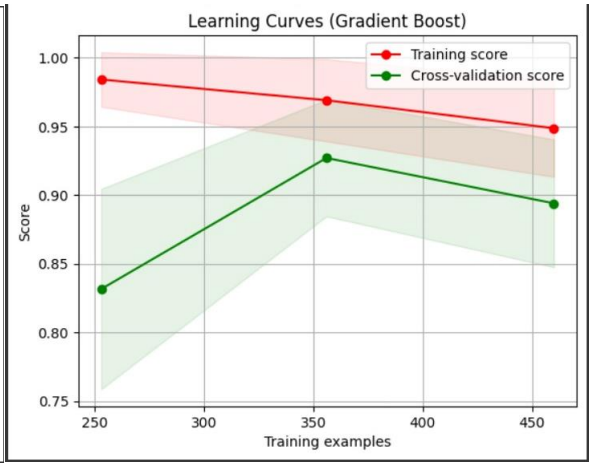

Fig. 16. ROC of Gradient Boost

2.6. Integration with Web Framework

To enhance practical applicability and ease of testing, we integrated the trained model with a user-friendly web interface. Flask is a web framework for Python that allows seamless integration of HTML and Python code. In this project we have used flask framework for the integration. The web interface contains various input fields like username length, full name length, description length, number of followers, number of people following them, number of posts, if profile picture exists, if external URL exists, if the account is private or public. The user has to fill the above input fields and get the results. This interface allows users to input profile information, which is then analysed for signs of being fake or genuine. Results, indicating whether the profile is classified as fake or genuine, are displayed through the web interface. This comprehensive methodology outlines the systematic approach employed in our study of fake profile detection, ensuring a rigorous and meticulous exploration of the subject matter.

2.7. UML Diagram

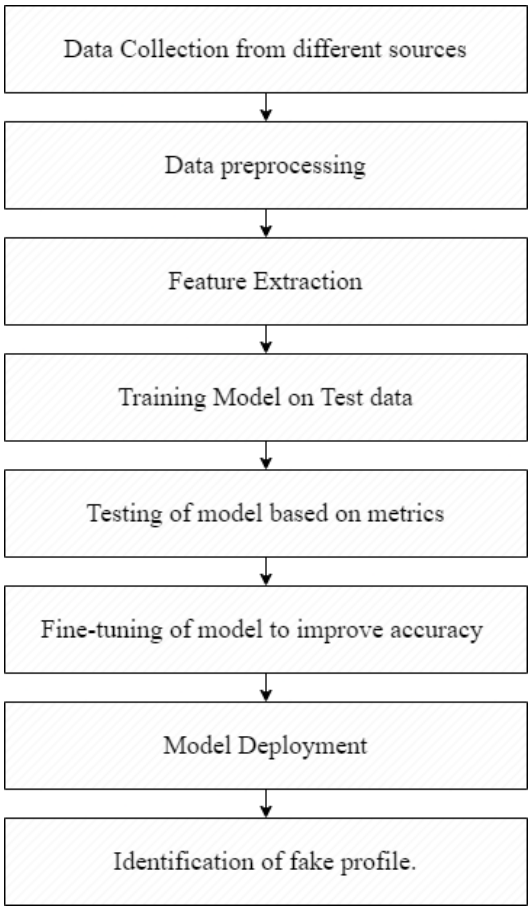

Fig. 17. UML Diagram of Fake Profile Detection

CHAPTER 3: DATA SET DESCRIPTION AND TECHNOLOGY STACK

3.1. Dataset Description

In this section, we delve into the intricacies of our dataset specifically designed for fake profile detection, compiled from Instagram sources. This amalgamated dataset comprises a plethora of profile attributes, ranging from profile names and description lengths to status and external URLs. Through meticulous curation, we ensured the dataset's relevance and comprehensiveness, meticulously addressing any overlaps to create a robust corpus for analysis and modeling in the realm of fake profile detection [13].

3.2. Dataset Details

The dataset consists of profiles sourced from Instagram, categorized into two primary groups: fake and real. For the purpose of training and testing machine learning models, the data distribution is as follows:

Table 1. Dataset Analysis

Class Label	Training Dataset	Testing Dataset
Real	288	60
Fake	288	60
Total	576	120

This split ensures a balanced representation of both fake and real profiles in both the training and testing phases, facilitating robust model evaluation and validation.

3.3. Hardware Requirements

For training of different algorithms, GPU is required. The GPU used in Google Colab is typically an Nvidia Tesla K80 or T4. These GPUs are well-suited for accelerating deep learning tasks and offer substantial computational power for training and inference tasks on large datasets.

3.4. Software Requirements

1. Python
2. Jupyter notebook
3. Machine Learning

Libraries:

1. sklearn
2. Seaborn
3. Matplotlib

Algorithms

1. Random Forest
2. Support Vector Machine
3. Logistic Regression
4. Gradient Boost Classifier
5. K-Nearest neighbor
6. Multi Layer Perceptron
7. Naive Bayes

CHAPTER 4: RESULTS AND ANALYSIS

4.1 Summary

The results show that using advanced deep learning and Machine Learning Models like Naive Bayes, K-Nearest Neighbor, Multi Layer Perceptron, Gradient Boost Classifier, Support Vector Machine, and Random Forest, can help identify Fake Profile with high accuracy. The improvements in accuracy were seen as we moved from simpler models like to more advanced ones, showing that these models are effective at handling the complexities of Fake Profile Detection as shown in below Table 2.

To evaluate the performance as quality of the model, different metrics are used, and these metrics are known as performance metrics or evaluation metrics. This performance metrics include accuracy, precision score, confusion matrix, F1 score, recall.

1. Accuracy: The accuracy metric is one of the simplest Classification metrics to implement, and it can be determined as the number of correct predictions to the total number of predictions.

$$\text{Accuracy} = (TP+TN)/(TP+TN+FP+FN)$$

2. Precision: The precision determines the proportion of positive prediction that was actually correct. It can be calculated as the True Positive or predictions that are actually true to the total positive predictions

$$\text{Precision} = TP/(TP+FP)$$

3. Recall: t can be calculated as True Positive or predictions that are actually true to the total number of positives, either correctly predicted as positive or incorrectly predicted as negative

$$\text{Recall} = TP/(TP+FN)$$

4. F1 Scores: F1 Score can be calculated as the harmonic mean of both precision and Recall, assigning equal weight to each of them.

$$\text{F1 Score} = 2* (\text{Precision} * \text{Recall})/(\text{Precision} + \text{Recall})$$

Table 2. Results for ML and DL Models

Sr.No	Model	Accuracy	Precision	Recall	F1 Score
1	Random Forest	94	91	91	91
2	Support Vector Machine	74	81	73	72
3	K-Nearest Neighbours	88	86	86	86
4	Naïve Bayes Classifier	70	67	62	58
5	Logistic Regression	91	92	92	92
6	Gradient Boost	89	93	93	92
7	Multi Layer Perceptron	96	90	90	90

CHAPTER 5: CONCLUSIONS AND FUTURE SCOPE

5.1. Conclusions

In summary, our research concentrated on detecting fake profiles specifically on Instagram, employing a range of machine learning and deep learning models. Through meticulous experimentation and evaluation, we gained valuable insights into the performance of various models in identifying fake profiles. Notably, our study showcased that training the model using a multi-layer perceptron significantly boosted accuracy to 96%, surpassing the base model's accuracy of 91%. This contribution adds to the ongoing efforts in detecting and mitigating fake profiles on Instagram, offering insights and methodologies for constructing effective fake profile detection systems. Looking ahead, further research into advanced deep learning architectures, fine-tuning hyperparameters, and leveraging specialized embeddings shows promise for advancing the field of fake profile detection and fostering respectful online interactions in multilingual communities.

5.2. Future Scope

Moving forward, there are several avenues for enhancing the effectiveness of our approach. Firstly, improving the accuracy of the created models remains a priority. This could involve refining the training process, exploring additional features, or fine-tuning hyperparameters to achieve higher precision and recall rates. By expanding the scope of our research, we can address the global challenge of online fake profile detection more comprehensively, fostering a safer and more inclusive digital environment.

References:

[1] R. Singh Boparai & R. Bhatia. Detection of fake profiles in online social networks-a survey. https://ssrn.com/abstract=4159087.

[2] A. Bhattacharya, R. Bathla, A. Rana & G. Arora. (2021). Application of machine learning techniques in detecting fake profiles on social media. In: 9th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions). DOI: 10.1109/ICRITO51393.2021.9596373.

[3] Elovici, Yuval, and Gilad Katz. "Method for detecting spammers and fake profiles in social networks." U.S. Patent No. 9,659,185. 23 May 2017

[4] T. Om Prathyusha, N. S. Kumar & E. V. Priya. (2021). Fake account detection using machine learning. Available at: www.ijcrt.org

[5] Gurajala, Supraja, et al. "Profile characteristics of fake Twitter accounts." Big Data & Society 3.2 (2016): 2053951716674236.

[6] Preethi Harris, "Fake Instagram Profile Identification and Classification using Machine Learning" 2021 2nd Global Conference for Advancement in Technology (GCAT) Bangalore, India. Oct 1-3, 2021.

[7] Jyothis Joseph, Vineetha S , "Fake Profile Detection in Online Social Networks using Machine Learning Models". DOI: 10.1109/RASSE60029.2023.10363482

[8] Dr.M.Kathiravan, Dr. Shaik Javed Parvez, "Analysis and Detection of Fake Profile Over Social Media using Machine Learning Techniques". DOI: 10.1109/ICSSIT55814.2023.10061020

[9] Ms.S.R. Ramya AP/CSE, Ms.R. Priyanka, STUDENT/CSE, "SVM Based Fake Account Sign-In Detection". 7th International Conference on Trends in Electronics and Informatics (ICOEI 2023)

[10] Kunal Umbrani, Deven Shah, "Fake Profile Detection using Machine Learning". DOI: 10.1109/CETSIS61505.2024.10459570

[11] Jai Gangan, Suprith KS, “Detection of Fake Twitter Accounts using Ensemble Learning Model”. Pimpri Chinchwad College of Engineering (PCCOE), Pune, India. Aug 18-19, 2023.

[12] Muraya, Catherine, Fredrick Awuor, and Benard Maake. "Fake Profile Identification on Online Social Networks." Journal of Language, Technology & Entrepreneurship in Africa 12, no. 2 (2021): 48-59

[13] https://www.kaggle.com/datasets/free4ever1/instagram-fake-spammer-genuine-accounts

[14] Muñoz, Samuel Delgado, and Edward Paul Guillén Pinto. "A dataset for the detection of fake profiles on social networking services." In 2020 International Conference on Computational Science and Computational Intelligence (CSCI), pp. 230-237. IEEE, 2020.

[15] Shinde, Shruti, and Sunil B. Mane. "Malicious profile detection on social media: A survey paper." In 2021 9th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions)(ICRITO), pp. 1-5. IEEE, 2021.

[16] Van Der Walt, Estée, and Jan Eloff. "Using machine learning to detect fake identities: bots vs humans." IEEE access 6 (2018): 6540-6549

[17] Bharti, Nasib Singh Gill, and Preeti Gulia. "Exploring machine learning techniques for fake profile detection in online social networks." Int. J. Electr. Comput. Eng. IJECE 13, no. 3 (2023): 2962.

[18] Bindu, A. Hima, Sheelam Sriya Reddy, Gurrala Sai Sri, and U. G. Students. "Fake Profile Identification in Social Network Using Machine Learning and NLP." (2023).

[19] Kaushik, Keshav, Akashdeep Bhardwaj, Manoj Kumar, Sachin Kumar Gupta, and Abhishek Gupta. "A novel machine learning-based framework for detecting fake Instagram profiles." Concurrency and Computation: Practice and Experience 34, no. 28 (2022): e7349.

[20] Ezarfelix, Juandreas, Nathannael Jeffrey, and Novita Sari. "A Systematic Literature Review: Instagram Fake Account Detection Based on Machine Learning." Engineering, MAthematics and Computer Science Journal (EMACS) 4, no. 1 (2022): 25-31.

[21] Voitovych, O., L. Kupershtein, and V. Holovenko. "Detection of Fake Accounts in Social Media." Кібербезпека: освіта, наука, техніка. Том 2,№ 18: 86-98. (2022).