Project Report

Detection of Fake Social Media Profiles Using Machine Learning Algorithms

Done by:

1. **Hassan Kaif M**
2. **Mohammed Muhsin A**

Project Title: Detection of Fake Social Media Profiles Using Machine Learning Algorithms.

Project Objective: Our objective is to develop a robust and scalable Machine Learning model to identify fake social media profiles with high accuracy.

Background: The prevalence of fake social media profiles disrupts online discourse and undermines trust in these platforms. These fake profiles can be used to spread misinformation, manipulate public opinion, and target users with malicious content. Fake profiles are easily identified: fake accounts frequently have a low number of followers, follow a large number of accounts, have fewer posts, and sometimes have longer bios with URLs, whereas real accounts have a more balanced follower and following count, varying numbers of posts, shorter bios, less likely to have URLs in bios, and usually have profile pictures.

Methodology: This project utilizes Ensemble Learning, which is a ML technique that combines multiple models to achieve better predictive performance than any single model as no one single model is always the most accurate. We have used Random Forest, Logistic Regression, and SVM models for this project. The model combination scheme that we have utilized is the Voting Classifier scheme. By training a diverse set of base learners on labelled datasets of real and fake profiles, our project aims to achieve a high degree of precision.

Project Impact: This project will contribute to a safer and more reliable social media environment by:

1. Reducing the effectiveness of fake profiles in spreading misinformation and manipulating online discourse.
2. Protecting users from malicious content and fraudulent activity.
3. Providing social media platforms with tools to identify fake accounts which can then be removed by these platforms.
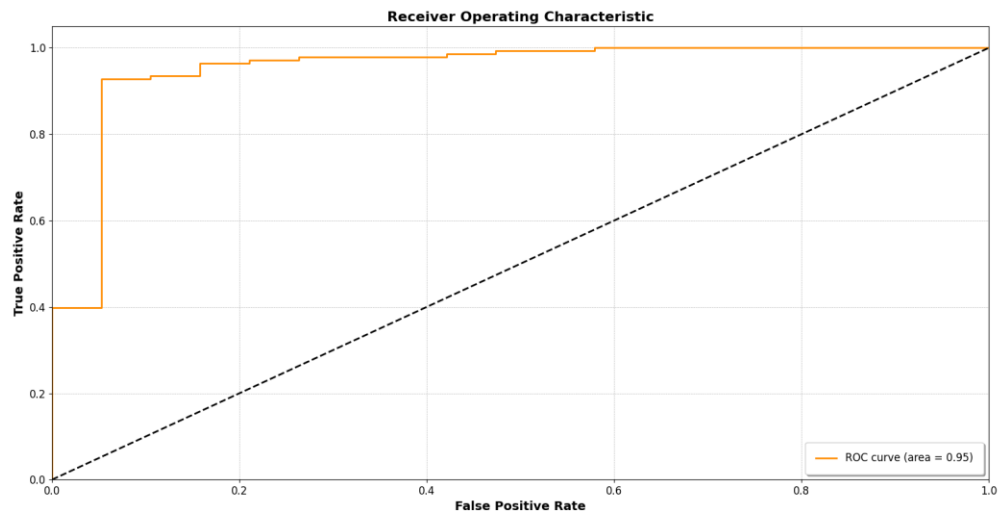
How it works: A dataset is used to develop and train the three models initially. We used a Kaggle dataset of fake Instagram profiles. The dataset that we have used can be found here: [https://www.kaggle.com/datasets/rezaunderfit/instagram-fake-and-real-accounts-dataset]. We will now use a testing dataset to test the models and determine the model's accuracy once they have been trained. When this is finished, we will use Ensemble Learning's Voting Classifier scheme to combine the three learning models. In our project, we have employed the Hard Voting Classifier approach. The class with the largest majority of votes, or the class with the best likelihood of being predicted by each classifier, is the predicted output class in hard voting. Now that the project has been tested by entering the required data, we will examine the results produced by the Hard Voting Classifier scheme, which combines the three learning models. We use a website that we have set up for the data entry and output procedures. HTML and CSS were used in the creation of the website. After that, the output is verified.

Working Diagram:
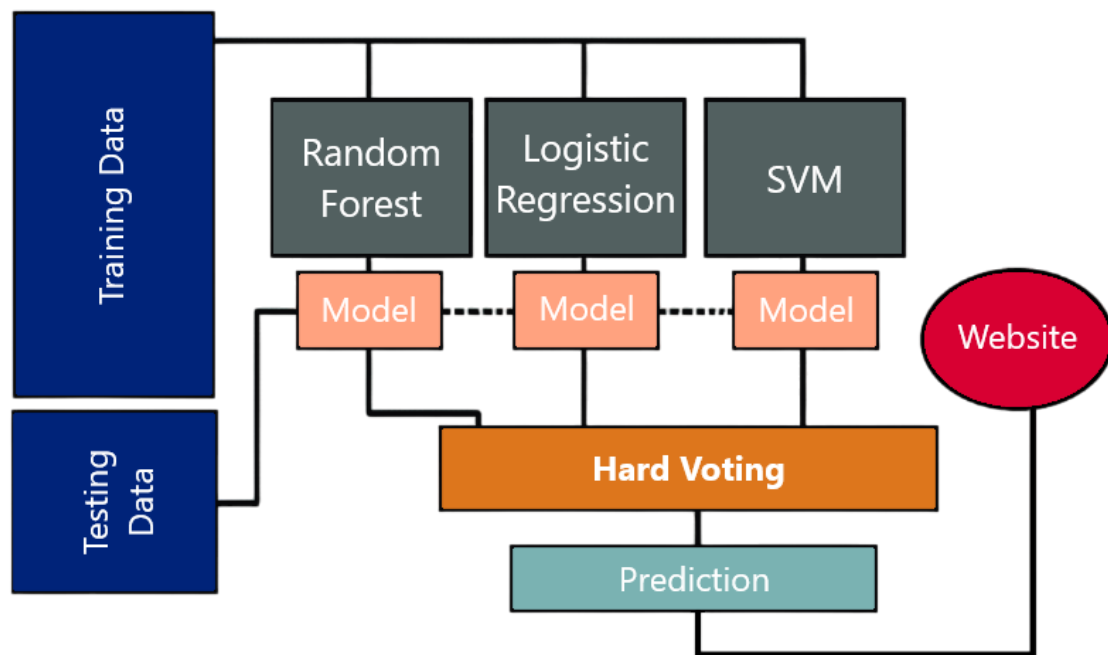
ROC Curve:

Precision-Recall Curve: