Gokaraju Rangaraju Institute of Engineering and Technology

(Autonomous)

Department of Information Technology

A Mini Project on

JOB SCAM ALERT

Guide

Dr. Y J Nagendra Kumar

Head of the Department

Team Members

(21241A12F6) J. Krishna Chaitanya

(21241A12D9) C. Abhinav Reddy

(21241A12G4) Vinod Pawar

Contents

- Abstract
- Aim and Scope
- Existing System
- Proposed System
- System Specifications
- References

Abstract

In response to the increasing prevalence of fraudulent job postings, we have developed "Job Scam Alert," an advanced system utilizing machine learning techniques to detect fake job advertisements. Our solution leverages Support Vector Machine (SVM) and Random Forest algorithms, achieving accuracy rates of 95.04101416853095% and 96.84936614466816%, respectively. "Job Scam Alert" effectively distinguishes between legitimate and scam job listings, providing a critical tool for job seekers to protect their personal information and make informed decisions in the digital job market. Through continuous research and refinement, we aim to enhance the system's accuracy and reliability, ensuring a safer online environment for all users.

Keywords: Job Scam Alert, Random Forest, SVM.

Domain: Machine learning, NLP.

Introduction

- Employment scam is one of the serious issues in recent times addressed in the domain of Online Recruitment Frauds (ORF).
- In recent days, many companies prefer to post their vacancies online so that these can be accessed easily and timely by the job-seekers.
- However, this intention may be one type of scam by the fraud people because they offer employment to job-seekers in terms of taking money from them.
- Fraudulent job advertisements can be posted against a reputed company for violating their credibility.

Aim

- The Aim is to develop "Job Scam Alert," a machine learning-based system using SVM and Random Forest algorithms to accurately identify and flag fraudulent job postings, thereby protecting job seekers from scams and enhancing online job market safety.

Scope

- The system applies advanced classification techniques to analyze job postings across various online platforms, aiming to improve detection accuracy and user trust in the digital job market, with ongoing enhancements for reliability and effectiveness.

Literature Survey

TITTLE	AUTHOR	PROS	CONS		
Machine Learning-Based Categorization Approaches to Avoid Fake Job Postings	Kumari and Sahani	- High accuracy (99%) - Uses various ML techniques	- Limited discussion on robustness and real-time detection	LinkedIn postings	99% accuracy
Predicting Fake Job Postings Using Naive Bayes and Decision Trees	Smith et al.	Real-time detection - Integrates topic modeling	- Computational complexity - Real-world effectiveness not detailed	Not specified	Not specified
Text Analysis for Detecting Fake Job Postings	Johnson et al.	- High accuracy (98.71%) - Uses Bi-LSTM	- Imbalanced dataset (5% fraudulent)	Not specified	98.71% accuracy, 0.91 ROC AUC score
Predicting Fake Job Postings Using Various Data Mining Techniques	Habiba and Islam	- High accuracy (98%) - Uses a range of classifiers	- Computational overhead - Challenges not extensively covered	EMSCAD (18,000 samples)	98% accuracy
Addressing Employment Fraud During the Pandemic	Kumari and Satya Kala	- Focus on pandemic impact - Uses NLP and ML	- Limited model performance details	Not specified	Not specified
Accurate Fake Job Detection Using Various Algorithms	Gulshan, Mukund, and Ajay A.	- High accuracy (98%) - Emphasizes continuous refinement	- Dataset specifics not detailed	EMSCAD (18,000 samples)	98% accuracy

Existing System

- Current job scam detection methods rely on basic keyword matching and heuristic rules, which often fall short in identifying sophisticated scams. These systems lack advanced machine learning capabilities, resulting in lower accuracy and higher false positives or negatives.

Proposed System

- "Job Scam Alert" utilizes advanced machine learning techniques, specifically Support Vector Machine (SVM) and Random Forest algorithms, to analyze job postings. This system achieves high accuracy in detecting fraudulent job listings by leveraging complex pattern recognition and data-driven insights, improving overall detection reliability and user protection.

System Specifications

Software

- ➢ Anaconda Navigator/Jupyter Notebook.
- ➢ Python 3.7 or later.
- ➢ Python libraries:
- pandas
- wordcloud
- numpy
- mathplot
- scikit-learn
- seaborn.

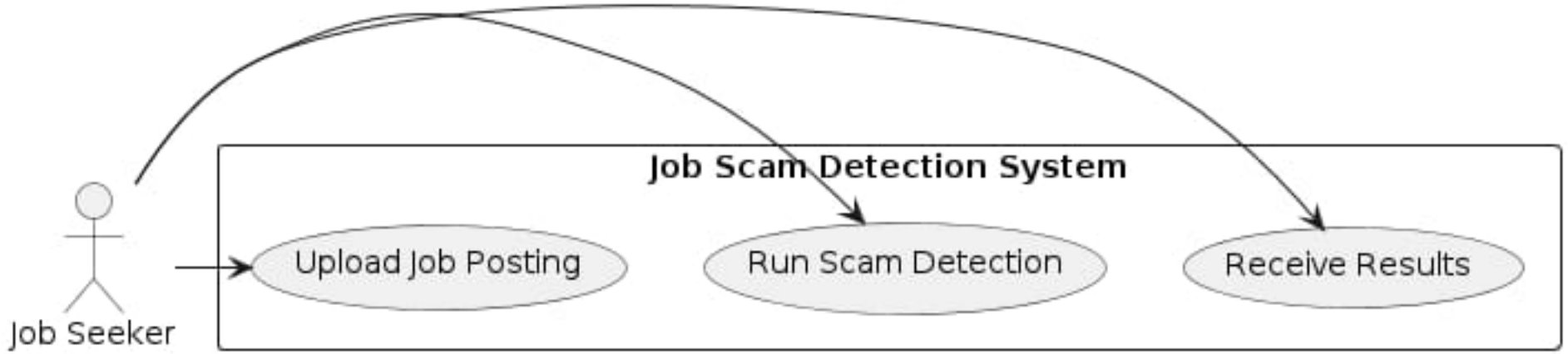
Hardware

- Processor –Intel core i5 and above
- Memory – 4GB RAM (Higher specs are recommended for high performance).
- Input devices – Keyboard,Mouse
- Internet

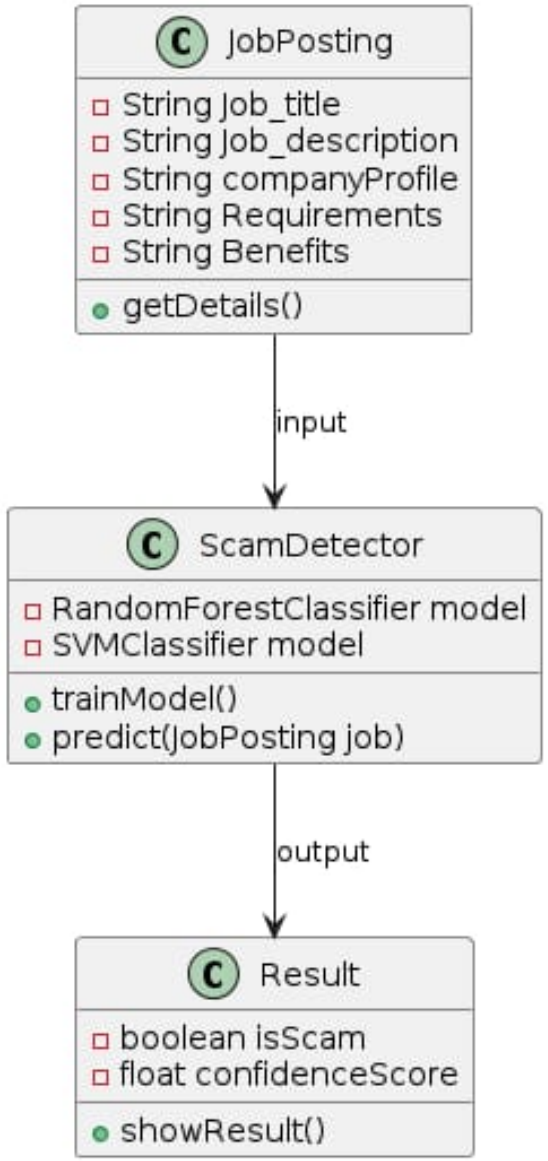
Dataset Description

Column	Description	Data Type	Valid Entries	Missing Entries	Unique Entries	Most Common Value
job_id	Unique Job ID	Integer	17,880	0	17,880	N/A
title	Job title	Text	17,880	0	11,200	"English Teacher Abroad" (2%)
location	Geographical location	Text	17,534	346	3,105	"GB, LND, London" (4%)
department	Corporate department	Text	6,333	11,547	1,337	"Sales" (3%)
salary_range	Indicative salary range	Text	2,868	15,012	874	"0-0" (1%)
company_profile	Company description	Text	14,572	3,308	1,709	"We help teachers get safe & secure jobs abroad :)" (4%)
description	Detailed job description	Text	17,880	0	14,800	Various
requirements	Job requirements	Text	15,186	2,694	12,000	"University degree required. TEFL / TESOL / CELTA or teaching experience preferred but not necessaryCanada/US passport holders only" (2%)
benefits	Job benefits	Text	10,674	7,206	6,207	"See job description" (4%)
telecommuting	Telecommuting position (0 or 1)	Boolean	17,880	0	2	0 (96%)
fraudulent	Target variable (0: Real, 1: Fake)	Boolean	17,880	0	2	0 (93%)

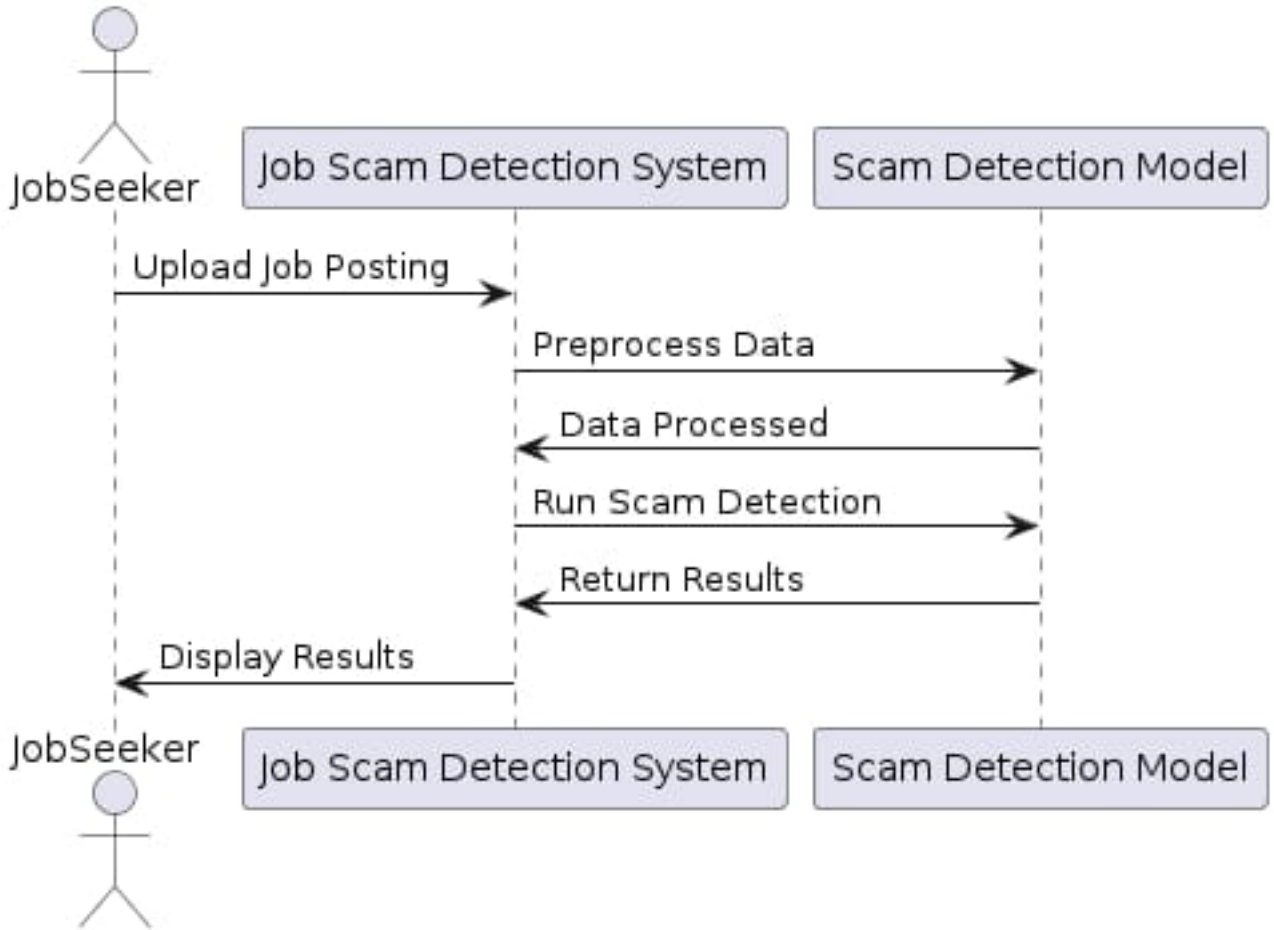
USECASE DIAGRAM



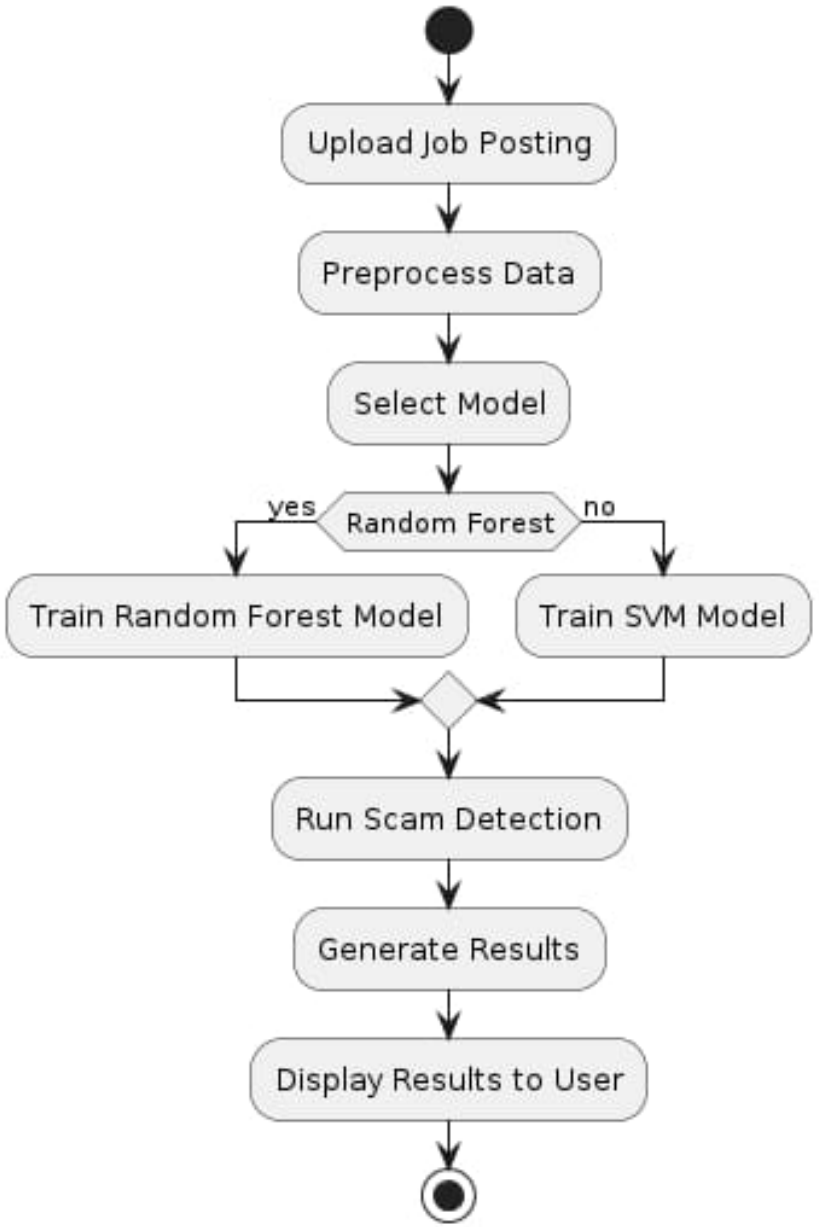
Class Diagram



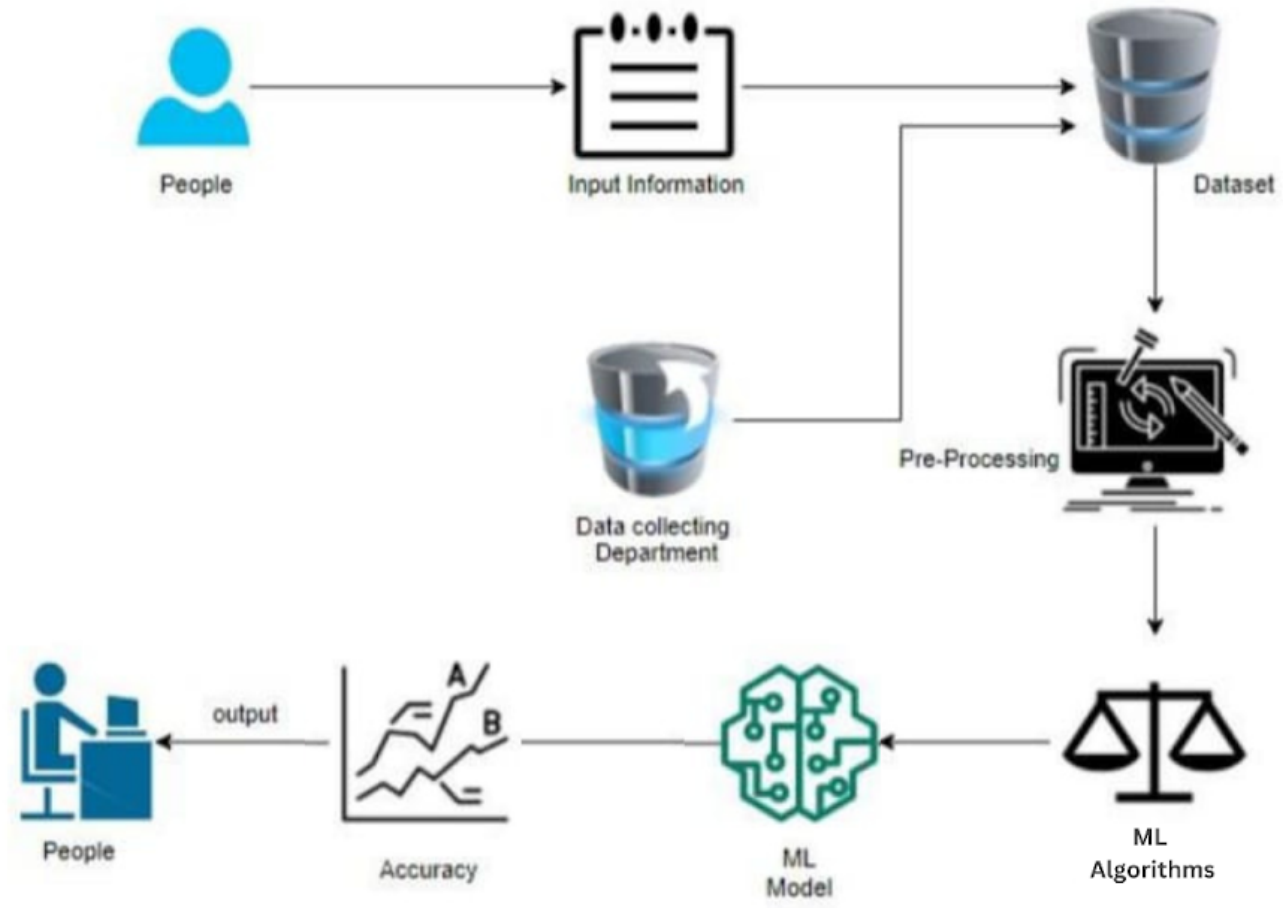
Sequence Diagram



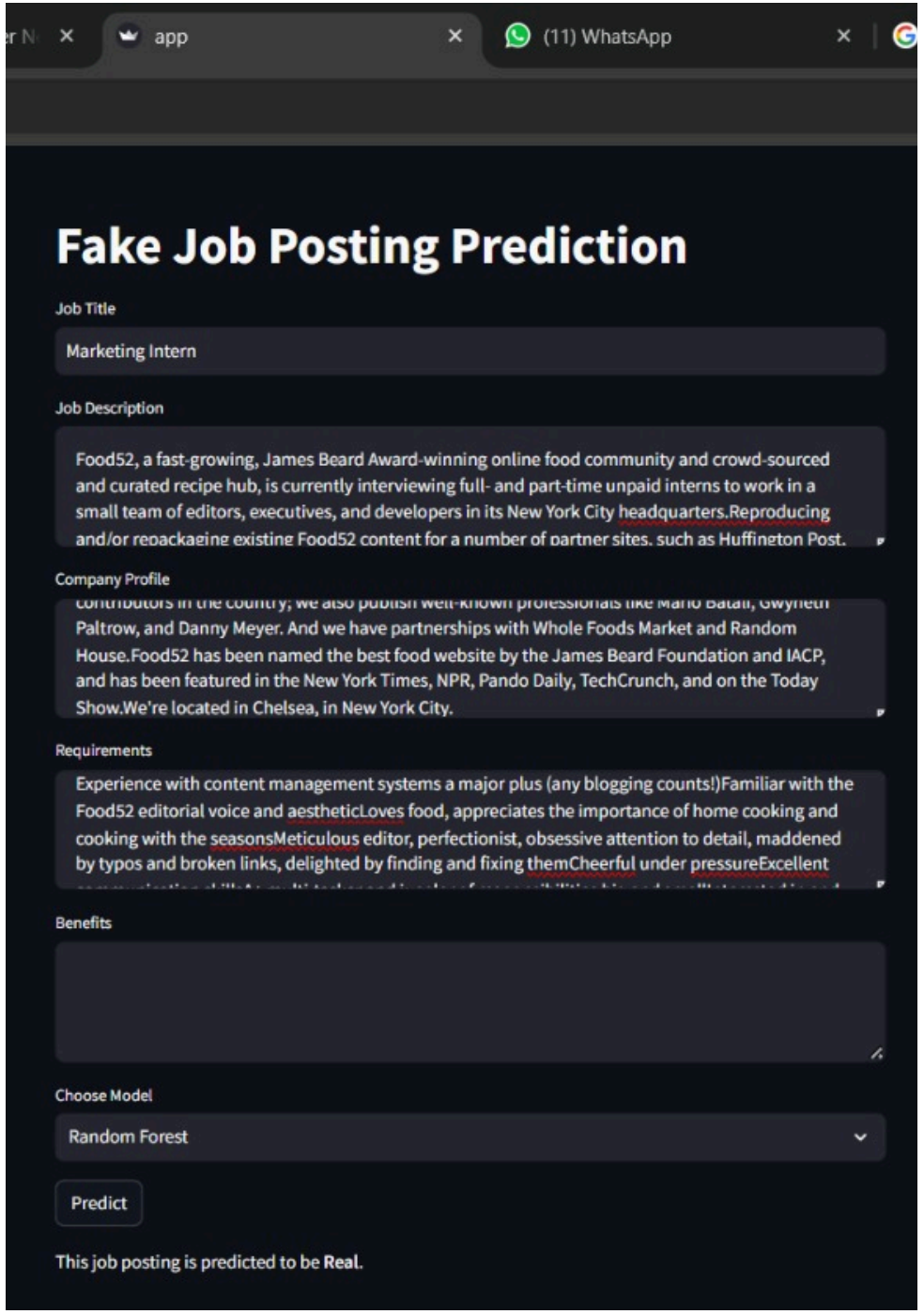
Activity Diagram



System Architecture



Final Output :



References

1. https://jespublication.com/uploads/2024-V15I40252.pdf
2. https://www.researchgate.net/publication/380074605_PREDICTING_FAKE_JOB_POSTS_USING_MACHINE_LEARNING_MODELS
3. https://arxiv.org/pdf/2304.02019
4. https://www.researchgate.net/publication/349884280_A_Comparative_Study_on_Fake_Job_Post_Prediction_Using_Different_Data_mining_Techniques
5. https://www.cnbc.com/2020/10/06/job-scams-have-increased-during-the-covid-19-crisis-how-to-one.html
6. https://www.ijfmr.com/papers/2024/1/13906.pdf