Project Name: TrustLens

Introduction:

- The rapid rise of AI and internet accessibility has led to a surge in sophisticated digital fraud, posing significant threats to individuals and society.

- This project focuses on two critical and growing areas of deception: **Deepfake Videos**, used for misinformation and impersonation, and **Fake Job Postings**, designed to steal personal data and money.

- Our goal is to develop "**TrustLens**," a single, user-friendly application and website that integrates detection models for both of these fraud types, providing a one-stop solution for digital safety.

Motivation:

- **Protecting Individuals:** Fake job scams lead to devastating financial loss and identity theft for job seekers. Deepfakes can be used for blackmail, defamation, and creating social unrest, causing severe emotional and reputational damage.

- **Eroding Trust:** The spread of deepfakes erodes trust in digital media and public institutions. Similarly, fake job postings undermine the credibility of online recruitment platforms and legitimate employers.

- **The Accessibility Gap:** While some detection tools exist, they are often separate, technically complex, or not easily accessible to the average internet user. There is a pressing need for a simple, reliable, and integrated tool.

Name : Sankalp Gupta   Roll No : 2201176   Supervisor : Dr. Angshuman Jana

Existing work / Literature review:

- **Deepfake Detection Apps:** Several specialized tools and online services exist (e.g., Deepware Scanner, DeepDetekt). However, they are single-purpose applications, forcing a user to seek out different tools for different problems. They are not part of an integrated fraud detection ecosystem.

- **Job Fraud Detection:** There are very few dedicated, user-facing applications where a person can independently verify a job post they received via email or a third-party site. Users mostly rely on manual checklists and intuition.

- **Integrated Solutions Survey:** A comprehensive review of the Google Play Store, Apple App Store, and web application repositories found **no publicly available application** that combines deepfake video analysis with fraudulent job posting detection in a single, unified interface.

Gaps in existing solutions/work:

- **Fragmented User Experience:** Users need to find, learn, and use multiple different websites or apps to check for different types of fraud. This is inefficient and inconvenient.

- **Lack of a Centralized User-Friendly Platform:** Most existing solutions are either academic prototypes or commercial products with limited features. There is a clear gap for a free, accessible, and easy-to-use platform for the general public.

- **Limited Scope:** Current tools are hyper-focused on one type of fraud. They fail to address the reality that a person susceptible to one type of online scam is a target for others. Our holistic approach provides broader protection.

Proposed Solution :

Deepfake Video Detection Module:

- **Video Processing:** Use **OpenCV** to automatically detect faces and extract sequential frames from the input video for analysis.
- **ML Model:** Implement a hybrid **CNN-LSTM** model. The CNN (**EfficientNet**) analyzes spatial artifacts in each frame, while the LSTM analyzes temporal inconsistencies to classify the video as 'Real' or 'Fake'.

Fake Job Posting Detection Module:

- **NLP & Feature Extraction:** Process the raw job description using a **BERT** model to convert the text into meaningful numerical vectors that capture its semantic context.
- **Classification Model:** A lightweight classifier (**Logistic Regression**) will take the BERT embeddings as input to accurately determine if the job posting is 'Legitimate' or 'Fraudulent'.

Application & Website Integration (End Goal):

- **Backend & API:** Deploy the trained models on a server using a **Flask REST API** to handle all analysis requests from the user interface.
- **Frontend Application:** Develop an intuitive web app (**ReactJs**) and mobile app (**React Native**) that will interact with the API, allowing users to easily upload a video or paste job text for an instant verdict.