Fake Instagram Account Detection Project Guide

Project Overview

Detect fake Instagram profiles especially those pretending to be girls to scam or mislead.

Tech used: Python, Instaloader, Scikit-learn, Regex, Gender Detector, Profile Pic Analysis

Step-by-Step Breakdown

1. Data Collection:

    - Use instaloader to extract public Instagram data like username, bio, followers, etc.

2. Feature Engineering:

    - Follower-following ratio, bio length, presence of profile pic, username pattern, engagement rate

3. Name-Based Gender Detection:

    - Use gender-guesser on names like 'priya', 'rahul' etc.

4. Detect AI-generated profile pics (optional):

    - Use CNN model or external services

5. Rule-Based Fake Detection:

    - Very low engagement + profile has feminine name = suspicious

6. Final Model:

    - Optional: Train Random Forest model with 500+ labeled data for better detection accuracy

Gender Detection via Name

Use gender-guesser Python library to predict name gender.

Example:

```
import gender_guesser.detector as gender
d = gender.Detector()
```

Fake Instagram Account Detection Project Guide

d.get_gender("priya") -> 'female'

Suspicious Username Patterns

Regex example:

If username contains 'sexy', 'babe', or long digit patterns mark suspicious.

e.g. 'priya_babe_9090', 'angel_hot_1001'

Final Classification Logic

If profile has:

- Followers < 200

- Followings > 1000

- Zero posts or bio with seductive emojis

- Detected as 'female' by name

Then it's very likely fake.

Tech Stack

- Python, Instaloader (for scraping)

- gender-guesser (for name detection)

- Regex (username pattern)

- Scikit-learn (ML)

- Optional Flask + React for web app interface