Fake Profile Detection On Matrimonial Sites Using Random Forest Classifier

Fr. Conceicao Rodrigues College Of Engineering,Bandra,Mumbai 50
Department Of Computer Engineering

Problem Statement

On a matrimonial portal, most of the users have a genuine intention of finding his/her desired life partner. However, due to various factors like monetary benefits and data crawling, it attracts few people with malicious intentions. Thus the existence of such fake profiles on the platform is highly undesirable to the business and other genuine users.

Methodology

The website has been made using HTML, CSS & JS and the backend has been implement using Flask. The machine learning model for fake profile detection is made using python. Random forest classifier algorithm is used as :--

a) Random forest algorithm has given the highest accuracy (0.9782430213464697), as compared to other two algorithms.

b) It is most accessible for tabular data.

Algorithms Used		
1. Support Vector Machine Supervised Machine Learning Algorithm	2. Artificial Neural Networks Supervised Machine Learning Algorithm	3. Random Forest Classifier Supervised Machine Learning Algorithm
Algorithm Chosen: Random Forest Classifier		

Table 1. Algorithm Used

Architecture

There are 5 modules used in our project:

- Registration
- Email-Verification
- Login
- Profile Creation
- Profile Detection (fake/real)

The module diagram of the entire system is shown in (Fig. 1).

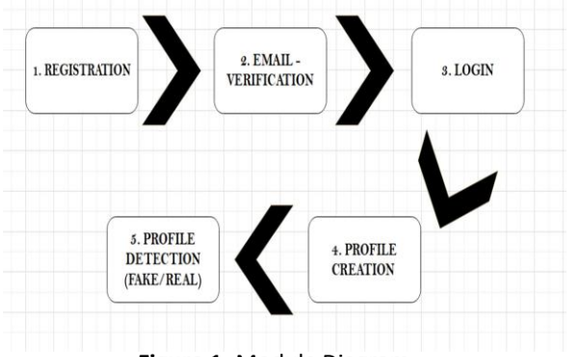

Figure 1. Module Diagram

Results

After the user fills in the required details, depending on the inputs provided, the profile is classified as Fake or Real .

In (Fig. 2) of the registration form, the user has entered the details which are all authentic, thus it is declared as a real profile:

1) Caste: Marwari

2) Degree: BE B.Tech (Engineering)

3) Employment Sector: MNC

4) Income: INR 20 lakhs to 25 lakhs

5) Mother Tongue: Rajasthani

6) Occupation: IT

7) Religion: Hindu

In (Fig. 3) of registration form, the user has entered the details which are all contradicting, thus it is declared as a fake profile:

1) Caste: 96K Kokanastha

2) Degree: 10th

3) Employment Sector: MNC

4) Income: INR Under 50 thousand

5) Mother Tongue: Garhwali

6)Occupation: Defence/Management/ Corporate Professionals

7) Religion: Buddhist

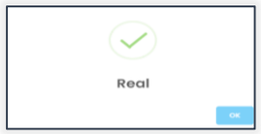

Figure 2. Real Profile

Figure 3. Fake Profile

Conclusion and Future Scope

Depending upon the details provided, our model is able to identify whether the profile is fake or real. Though the project initially focuses on fake profile detection, its scope can be extended to various features. The UI/UX of existing websites can be made more interactive . Other security facilities such as OCR can be incorporated.

Guide & Team Members

Supervisor :- Prof. Jagruti Nagaonkar

1. Riddhi Dinesh Oza
2. Ayukshi Rane
3.Godsee Almeida
4. Shaun Dabre