Identification of fake profiles using ANN

I. Introduction

Social Networking sites acts as a media of communication within people. Users use these sites for sharing of information that may be personal, economic, educational, business etc.,. These are also used for sharing of photos, videos and any day to day activities. However, some users aren't using for good objective. They create fake accounts with fake identity so we call them as an Attacker. They use these fake accounts and share fake news with affects the users. So. Identification of fake accounts is the major task for social sites. There are number of techniques for identifying fake accounts in the field of Machine Learning such as Neural Network, NLP and Classification. So we are going to identify the fake profiles with the help of Neural Network. Neural Network takes decision like a human brain as it consists of many interconnected processing elements.

II. Literature Review

Gayathri A, Radhika S, Mrs. Jayalakshmi S.L , have proposed a technique for detecting fake accounts in media application using Machine Learning. This paper presents some filtering algorithms that rely on classification to decide whether the profile is fake or not. This framework uses classification techniques like Support Vector Machine and Random Forest to classify whether the profile is genuine or fake. Data set of both fake and genuine profiles with various attributes like number of friends, followers are taken. The dataset is divided into training dataset and test dataset. They used a publicly available dataset of 1337 fake users and 1481 genuine users consisting of various attributes including number of friends, followers, status count, languages known etc., By using Random Forest 91% accuracy has been obtained and by using SVM 90.01% has obtained.

Snehal Bhambar, Kanchan Khairnar, Yogita Nikam, Harshali Shelar, Y.K. Desai , have proposed a technique for identification of fake accounts. This paper shows some of the mining tools which allows quick user interaction with a simple tool for the identification of fake accounts. The techniques used in this paper for detecting fake accounts are Neural Network(NN), Naive Bayes, Markov Model and Bayesian Network. Neural Networks take decisions like a human brain. SVM is used for classification. Naive Bayes classifier is used to predict the probability whether the variable belongs to a particular class. By using survey method the dataset of the Facebook or Twitter is collected. K-mediod clustering is used to increase the accuracy and reduce the time complexity of the algorithm.

Z. Halim et al. , have proposed a technique for identification of fake accounts. This paper shows spatio-temporal mining on social network to determine circle of customers concerned in malicious events with the support of latent semantic analysis. The techniques used in this paper for detecting fake accounts are, Natural language Processing (NLP) , Support Vector Machine (SVM) and Naive Bayes. An SVM classifies information by means of finding the exceptional hyperplane that separates all information facts of 1 type from those of the other classification. The best hyperplane for an SVM method that the one with the biggest line between the two classes. Naive Bayes is a probabilistic classifier, which means it predicts on the basis of the probability of an object. By using these By using these techniques, we can easily detect the fake profiles from the social network sites.

Sarode and Mishra , proposed a different approach which is a sequence of steps to detect fake profiles. The algorithm used is-first, the data is in JSON format, which is further parsed to a structured format (CSV) that is easier readable by machine learning techniques. These comma separated values will later make the classifier more efficient. The authors tried unsupervised and also supervised machine learning techniques. In this case, supervised machine learning techniques had a higher accuracy rate of almost 98%. Each equation at each neuron (node) is put through a Sigmoid function. We use a training data set by Facebook or other social networks. This would allow the presented deep learning algorithm to learn the patterns of bot behavior by backpropagation, minimizing the final cost function and adjusting each neuron's weight and bias.

Liang Wu, Fred Morstatter, Kathleen M. Carley and Huan Liu , have proposed a technique for Misinformation in Social Media. In this paper, we aim to consolidate the observations, and investigate how an optimal method can be selected. The techniques used in this paper for identifying fake profiles are Classification methods and Supervised Learning methods for detecting misinformation. Supervised learning methods have been studied to detect misinformation. They usually collect posts and their labels and then train a text classifier based on the collected content and labels. The key terms related to misinformation are disinformation, fake news, rumour, urban legend, spam, troll. And also in this paper Classification methods are also used.

III. Proposed System

In proposed system, we are using Artificial Neural Networks for the identification of fake profiles. ANN algorithm will be trained with all previous users fake and genuine account data and then whenever we gave new test data then that ANN train model will be applied on new test data to identify whether given new account details are from genuine or fake users. All fake users intention is to send friend request to all the users. After sending the request they demand money the users might not know that he is fake so they unknowingly sends money to them they came to know it is fake after few days but it will be too late. Not only demanding money may also do some illegal activities by asking photos , videos etc.,. To avoid such type of activites we are proposing a website where we can

check whether the request from that account is fake or real. By this most of the illegal activities can be avoided and users account can be secured. We are using ANN algorithm to identify the fake profiles in the social media. We are using this for one of the social networks it is instagram. We are introducing website to check whether the account is fake or real.

IV. Methodology

The identification of fake profiles mainly consists of three stages. At the first stage we need to design the model which is used to identify the profile whether it is fake or real. Second stage is creating a website which consists of two modules one is admin and second one is for users.

Admin:

Admin can change the model at any time according to the performance of the models. He/She can also update the dataset as per requirements. If any new modifications are required in the website then admin can do that changes these are the main responsibility of admin it is to maintain the website.

Users:

Users will create an account in it and enter the account details. All the required attributes will be mentioned in the website so the user will enter the required details and identify whether it is fake or real.

At the third stage we need to build an interface in order to interact with the model. Once the interface is built then ANN model will be used backend of the website. The architecture of the ANN is as shown below

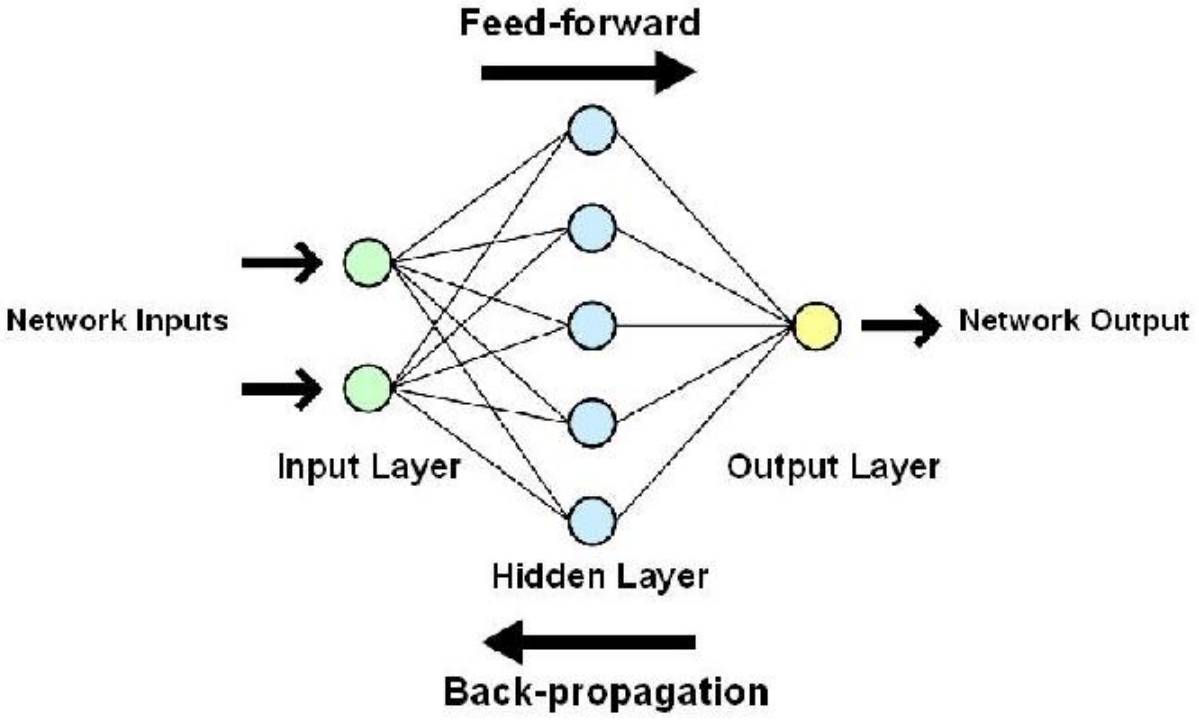

Fig.1: ANN Architecture

The above shown is the architecture of internal working of an ANN.It is to be trained with the training dataset and the weights, bias values are adjusted to get the desired output. The adjustment of the weights and bias values are done by using feed forward and back propagation. In feed forward, initially we use some random weights ranging between -0.5 to 0.5 and a bias value is taken at random. We calculate the output of the network, if the output is not matched with the original output then we calculate the error in the network. Once the error is calculated we back propagate the network and changes the weights and bias value in the network. This process is continued until the output is matched. Once the model is trained we test the model with the test dataset. Then the model is saved. The entire system architecture of this proposed work is as shown in below figure

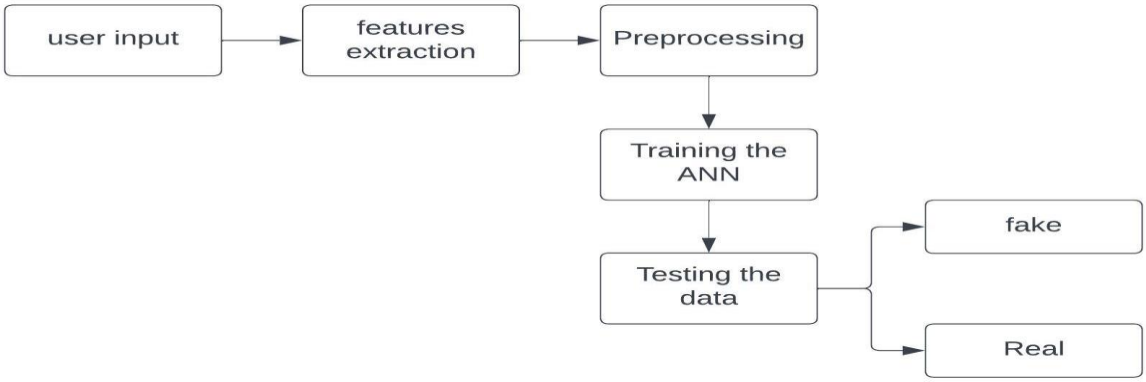

Fig.2: System Architecture

The above is the design of our proposed work it takes input from user and the model will extract the features after extraction preprocessing will be done. The trained model will be executed and it gives the output whether it is fake or real.

Conculsion:

At present every person has accounts in social networks and everyone is willing to become friends in social networks. Due to increase in demand to the social networks lots of people are trying to do frauds like creating fake accounts for passing misinformation and some others are doing cyber-crime. The model presented in this project will ask for user details to login. After login the user needs to enter the account details, which is fake in user's perspective this can be confirmed by the model which we proposed. This website will help to avoid many illegal activities by identifying whether the given account is fake or real.