International Journal of Innovations & Advancement in Computer Science
IJIACS
ISSN 2347 – 8616
Volume 5, Issue 11
November 2016

Social Networks Classification using DBN Neural Network based on Genetic Algorithm

EsmaelInanlo
Department of Computer Engineering
Islamic Azad University
Iran, BuinZahra

Javad MohammadZadeh
Department of Computer Engineering
Islamic Azad University
Iran, Karaj

Abstract-todays, social networks are very ubiquitous in all parts of the world and there are many users which must be classified in proper and different groups in order to convenient relations. Artificial neural network can be applied for this problem. The main idea of this paper is using from DBN (Deep Belief Neural network) because it has some advantages relative to other neural networks. Training of neural network is an NP_hard problem whichcannotbesolved in reasonable time order. Genetic algorithm can be applied for NP problems such as neural network training problem. In this paper, genetic algorithm has been applied for training of neural network and its convergence and stability had been investigated. Three neural networks are considered and compared with each other. The results showed that DBN neural network has proper performance relative to other compared ones.

Key words: social networks, artificial neural networks, genetic algorithm and classification

i. INTRODUCTION

Social network is a network in which there are some actors whom have some relations with together. These actors can be considered as individuals and organizations. Social network growing andmanynumbers of users in these networks caused that these networks be interesting field in science. Many academic fields such as social psychology, graph theory, sociology and statistics merged in social networks.One of the most challenging problems is clustering of peoples or organizations in social networks. Clustering means is that peoples or organizations must be placed in different groups based on some characteristics like roommate, skill, workmate, hobby and etc. An example of clustering is shown in fig.1. Left side FIG shows a social network without clustering and right side FIG shows a social network with clustering. In FIG.1, every node can be considered as individual or organization in which similar colors express individuals or organizations with similar interesting.

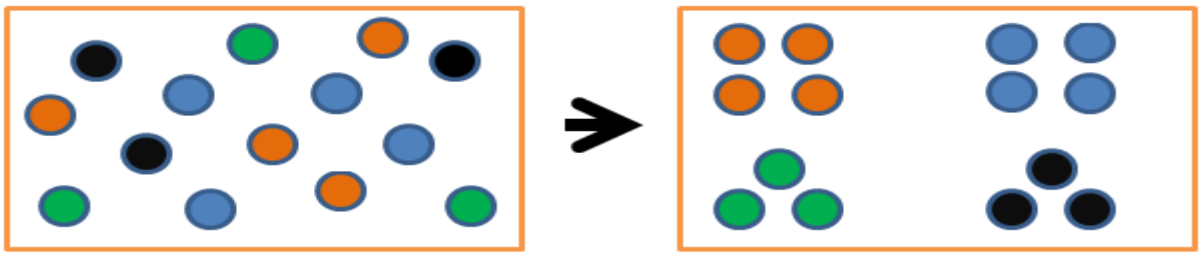

Figure.1: an example of clustering in social networks

Many approaches have been suggested to taxonomy that one of them is artificial neural networks (ANN). Neural network learning is an NP_hard problem which needs more and more execution time. Meta heuristic algorithm can be applied to these problems. We can consider the followings as related works:

Clustering of social networks with considering nodes' features and with using an online sequential method have been done. Not only did yungjiaoconsider the nodes features but also considered the interaction between linked nodes [1]. Rodrigo and et al introduced a new algorithm, which is based on lexicon coefficient classification, in order to classification of users' accounts in social networks [2].Social network context classification is a challenging problem. Susomrit done this taxonomy using a new algorithm and compared it with other algorithms [3]. In another recently done work, the taxonomy of social networks has been

done with considering frauds and money issues. This work's case study was europe [4]. Social network is converted to the point in which there is huge number of information. Many investigations have shown that this platform of information includes personal sensitive data, so data classification is crucial [5]. It is assumed that users are cooperative in social ad hoc social networks but selfishness is misbehavior in it. Selfishness detecting using ad hoc networks classification has been perused [6]. Users can be classified as confidence and none confidence. Artificial neural network can be applied for this problem. Also the trustworthiness is variable from one user to another, so they must be classified in different confidence groups [7]. We can find other useful information around social network areas and some novel ideas in [8][9][10].

ii. Artificial Neural Networks or ANN

ANNs are inspired by creature neural networks. There are neurons which connected to other neurons using connectors in ANN. Here is a sample of ANN which consists of input layer (red color nodes), one hidden layer (blue color nodes) and output layer (green color nodes). Number of input layer's nodes depends on problem nature. For social networks classification problem, number of individuals or organizations' features determines it. Number of hidden layer's nodesand number of hidden layers is determined by neural network designer.Also,number of output layer's nodes depends on nature of problem. Neural network training is specified by assigning values to its edges. It is an NP_hard problem. The training is acceptable if the goal of neural network, such as prediction, classification, etc, is satisfied.

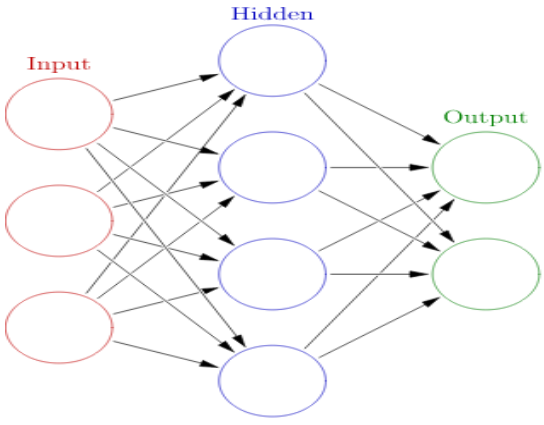

Figure.2: a sample of ANN [10]

iii. Proposed algorithm

We used from DBN artificial neural network, a full connected neural network, for clustering of users in social networks because:

a) Its training speed is acceptable even though we enhance number of hidden layers.

b) It does not catch in local optimal.

As mentioned before it, ANN training is an NP_hard problem so usingfrommeta heuristic algorithms is inevitable. Genetic algorithm is applied for training in DBN. A chromosome consists of gens and is a solution for the problem. Number of gens is equal with number of edges in ANN. We consider two layers for this problem except input and output layers. There are 10 attributes like university name, age, name, etcin input layer.We must considerneurons' values as numeric values. Therefor we assign numeric values for every member of groups. For example,azad university is presented by "1", Tehran university is presented by "2" and others are respectively represented. One of the other issues in neural network training is that the ranges of some values are higher than others. For example, range of age can be between 16 and 40 whereas the range of height can be between 150 and 200. As it is obvious, height attribute will have more effect on ANN than age will have. We use normalization in order to address this issue (equ.1).

$$y_{ij} = Y_{i\,min} + \frac{(y_{i,max} - y_{i,min}).(x_{ij} - x_{i,min})}{(x_{i,max} - x_{i},min)} \quad (1)$$

Y_{imin} and y_{imax} are the new range's minimum and maximum values. X_{ij} is a value that we want to normal it. And finally, $x_{i,max}$ and x_{imin} are the minimum and maximum of value of a group which must be normalized. Output layer has only one node and its value indicates the taxonomy of the user whose attributes had been given to input layer. The first population of chromosomes will be created after normalization step. One point crossover and random mutation operations are applied with 30% rate. Merit of a chromosome is variable from other chromosomes. We need a criterion for measuring chromosome's merit. This merit is known as fitness function. Eq.2 indicates how fitness value is calculated for a chromosome.

Fit =num - fnum(2)

Num and fnum are sequentially number of individuals which must be classified and number of false classified users. Higher is fitness value, more

appropriate is chromosome. We also use roulette selection operator in order to create new generation.

iv. Experimental results

Besides DBN neural network, one layer perceptron neural network and two layers perceptron neural network were also implemented on a system with 2.2GH of dual core CPU, 1GB of RAM and win7 operating system. Mentioned artificial neural networkswere implemented in MATLAB programing language. Two different datasets were acquired from facebook social network and telegram social networkandwere considered as inputs. Every dataset includes four different groups.We used genetic algorithm in order to trainingmentioned ANNs. Convergence and stability are the criteria that we can use to compare evolutionary algorithms [11]. In this paper, we examined the convergence and stability of genetic algorithm on mentioned artificial neural networks.Fig.3 shows convergence of ANNs and fig.4 shows stability of ANNs for the dataset of facebook social network.Also fig.5 and fig.6 show the convergence and stability of ANNs for the dataset of telegram.

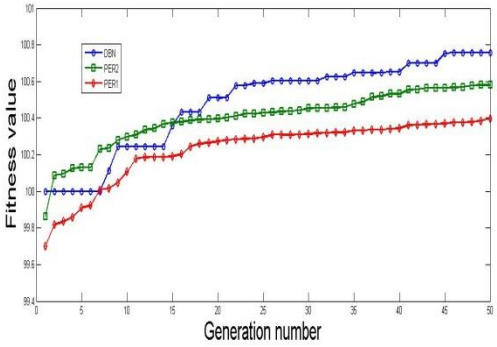

Figure.3: convergence of ANNs on facebook dataset

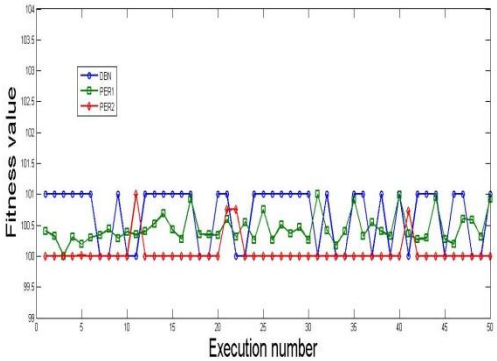

Figure.4: stability of ANNs on facebook dataset

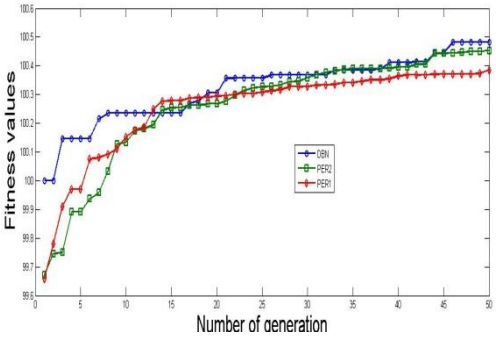

Figure.5: convergence of ANNs on telegram dataset

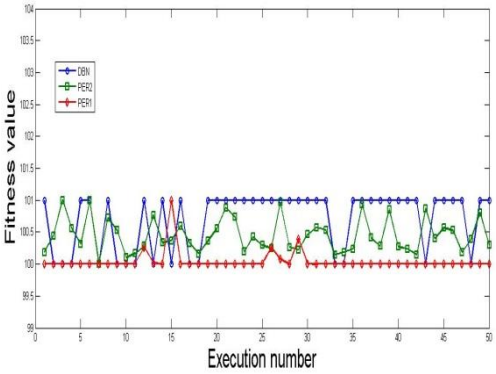

Figure.6: stability of ANNs on telegram dataset

DBN artificial network (DBN), two layers perceptron artificial (PER2) and one layer perceptron (PER1) had been sequentially showed with blue, green and red colors. All of them were trained using genetic algorithm. Convergence means is that the algorithm will close to optimal answer if number of generations or dedicated time is enhanced.For the facebook dataset, we can see that DBN has proper convergence than others have although it is weaker than PER2 until 14th generation. Also PER1 convergence is acceptable but it places under DBN and PER2. Other comparison criterion of the evolutionary algorithms is stability. Stability is referred to fluctuation of produced results on different individual executions. It is convenient that the produced results be close and similar to together on distinct executions. Stability was considered over 50 separate executions. Fig.4 presents ANNs' stability for the dataset obtained from facebook. Although PER1 has the least fluctuation, its mean value is lower than two other ANNs. Results of PER2 are weaker than PER1 in 4 executions. PER2 fluctuation is very slow and its results are close together. Like PER1 and PER2, DBN has acceptable stability but the mean value of its executions is higher than PER1

and PER2. We examined convergence and stability of genetic algorithm on telegram dataset like facebook dataset. From convergence view point we observe that all ANNs have appropriate convergence and have close competition, but DBN's convergence is better than two others are. The stability of ANNs on telegram dataset is similar to the stability of then on facebook dataset and we can apply same expression stated for fig.4.

v. Conclusion and feature works

We applied genetic algorithm to three types of artificial neural networks in order to classification of social networks. We examined genetic algorithm's convergence and stability on training on them. The results showed that genetic algorithm applied to the DBN is the more performance than other are. We can consider the followings as feature works:

a) Using other revolutionary algorithm for training of ANN and comparing them with the results obtained from this paper.

b) Using other artificial neural networks for classification of social networks and comparing them with the results of this paper.

c) *Applying DBN to other classification problems.*

REFERENCES

[1]Yongjiao Sun, Ye Yuan, Guoren Wang, An on-line sequential learning method in social networks for node classification, Neurocomputing, Volume 149, Part A, 3 February 2015, Pages 207-214, ISSN 0925-2312.

[2] Rodrigo Augusto Igawa, SylvioBarbonJr, Kátia Cristina Silva Paulo, GuilhermeSakaji Kido, Rodrigo Capobianco Guido, Mario LemesProençaJúnior, Ivan Nunes da Silva, Account classification in online social networks with LBCA and wavelets, Information Sciences, Volume 332, 1 March 2016, Pages 72-83.

[3] P. Susomrith and Y. Suseno, Chapter 12 - Social Capital and the Social Context of Business Networks: The Case of Thailand, In Business Networks in East Asian Capitalisms, edited by Jane Nolan, Chris Rowley and Malcolm Warner, Elsevier, 2017, Pages 269-288.

[4] RongRong, Daniel Houser, AnoviaYifan Dai, Money or friends: Social identity and deception in networks, European Economic Review, Available online 19 April 2016

[5] Alexandre Viejo, David Sánchez, Enforcing transparent access to private content in social networks by means of automatic sanitization, Expert Systems with Applications, Volume 62, 15 November 2016, Pages 148-160.

[6] Ahmedin Mohammed Ahmed, Xiangjie Kong, Li Liu, Feng Xia, SaeidAbolfazli, ZohrehSanaei, AmrTolba, BoDMaS: Bio-inspired Selfishness Detection and Mitigation in Data Management for Ad-hoc Social Networks, Ad Hoc Networks, Available online 14 September 2016.

[7] EbinDeni Raj, L.D. DhineshBabu, An enhanced trust prediction strategy for online social networks using probabilistic reputation features, Neurocomputing, Available online 22 September 2016.

[8] EbinDeni Raj, L.D. DhineshBabu, A fuzzy adaptive resonance theory inspired overlapping community detection method for online social networks, Knowledge-Based Systems, Available online 19 September 2016.

[9] Claudine Gravel-Miguel, Using Species Distribution Modeling to contextualize Lower Magdalenian social networks visible through portable art stylistic similarities in the Cantabrian region (Spain), Quaternary International, Volume 412, Part A, 15 August 2016, Pages 112-123.

[10] Olga Peled, Michael Fire, LiorRokach, Yuval Elovici, Matching entities across online social networks, Neurocomputing, Volume 210, 19 October 2016, Pages 91-106.

[11] YosefMasoudi-Sobhanzadeh, HabibMotieghader, World Competitive Contests (WCC) algorithm: A novel intelligent optimization algorithm for biological and non-biological problems, Informatics in Medicine Unlocked, Volume 3, 2016, Pages 15-28.