



Operations Research Society in
Bosnia and Herzegovina

Southeast Europe Journal of Soft Computing
Available online: <http://scjournal.ius.edu.ba>



IUS Soft Computing
Research Group

Authorship Authentication of Short Messages from Social Networks Using Recurrent Artificial Neural Networks

N. M. Demir

Faculty of Engineering and Natural Sciences,
International University of Sarajevo International University of Sarajevo,
Hrasnicka Cesta 15, Ilidža 71210 Sarajevo,
Bosnia and Herzegovina
ndemir@ius.edu.ba

Article Info

Article history:

Article received on 17 June 2018
Received in revised form 10 August
2018

Keywords:

Authorship Authentication; short
messages; committee machines;
recurrent neural network

Abstract

Dataset consists of 17000 tweets collected from Twitter, as 500 tweets for each of 34 authors that meet certain criteria. Raw data is collected by using the software Nvivo. The collected raw data is preprocessed to extract frequencies of 200 features. In the data analysis 128 of features are eliminated since they are rare in tweets. As a progressive presentation, five – ten – fifteen – twenty - thirty and thirty four of these 34 authors are selected each time. Since recurrent artificial neural networks are more stable and iterations converge more quickly, in this work this architecture is preferred. In general, ANNs are more successful in distinguishing two classes, therefore for N authors, $N \times N$ neural networks are trained for pair wise classification. These $N \times N$ experts then organized as N special teams (CANNT) to aggregate decisions of these $N \times N$ experts. Number of authors is seen not so effective on the accuracy of the authentication, and around 80% accuracy is achieved for any number of authors.

1. INTRODUCTION

The Internet has dramatically become significant. Social Networks have taken interest of billions and their effect grows each day. Users reach the others, share their opinions and transmit the information. Online networks like Twitter and Facebook serve as virtual environment with simplicity and became rich and easy content platforms that provide

knowledge. Nonetheless, there are several security issues that occur with the wide usage of these sites. It can be considered that these sites have trustable environment but they are accessible to virtual attacks. Detecting fake and compromised account, and distinguishing them are the main problems in authorship authentication for social networks.

This work aims the study of developing a system which is able to operate for finding the author of anonymous messages by providing to the system posts written by a list

of suspected users on social media and choosing matched author.

Related studies investigated mostly focusing on longer text documents rather than what is intended to do by this research (Can,2012). This study is important by combining stylometry which is more than a century-old science with current computational capacity for short text messages. The stylometry regarding text classification of short social network messages, appropriate methods applied in relevant and contemporary research were investigated as the base of this study.

Stylometry, also known as authorship analysis purposes to determine the original author of a given text which studies linguistic style. The methods of it have been primarily applied to analyze letters and literary works such as Federal Papers (Hamilton, et. al., 2008). The analysis in the vocabulary of an author and the use-frequency of words in it are known as a general method in stylometry which is later compared with the vocabulary of another author. The specific analysis of the use-frequency of function words including numerals, pronouns, prepositions, auxiliary verbs, and conjunctions is also possible with it. The analysis of average sentence length or the use of very unusual words is another method applied with stylometry for comparing texts.

There are three main perspectives regarding today's applications. These are authorship attribution, authorship verification, and authorship profiling. Authorship attribution aims to determine a probable author from a multitude of several other authors. On the other hand, authorship verification finds if an author's linguistic style matches to another linguistic style of the author. Authorship profiling has the purpose of determining attributes which are likely to reveal an anonymous authors origin, age, gender, and so on. This work focuses on the first perspective, i.e. authorship attribution.

The detection of the authorship for a document which is fewer than 1000 words was thought to be difficult in the time of the early 19th century. In the early 21st century, the number decreased and the determination of the authorship of a document with 250 words was thought to be possible. There is also a need for decreasing this limit because of spreading usage of many shorter communication tools such as Twitter, Facebook etc.

There are differences between authorship attribution of online documents and the authorship attribution of traditional work. This occurs in two ways. The first is that the online documents or text collection are frequently informal and unstructured which are not necessarily grammatically correct as a comparison to literature texts. The second is that the quantity of authorship disputes regarding a single online document is much more as a comparison to traditional published documents. In this situation, the scarcity of standardized data to test the

accuracy of results underlies as the reason that is one of the challenges of authorship attribution.

For the researchers, the increasing of the popularity of social media has made it easier directing the focus on authorship attribution in micro-blogs. Various studies have been published as a respect to the use of authorship analysis in social network recently.

The problem of authorship attribution for an online social network Twitter is studied in this work. Twitter has had an increase with its popularity recently by reporting to have over 500 million user base that share almost the same quantity of messages daily which is called as tweets (internet live stats). Twitter differs from other social networks in terms of publishing limitation. Users are able to publish only 140 characters for each tweet.

Various classification methods are implementable to the authorship attribution problem. An important transition from statistical methods into machine learning based approaches is demonstrated by the authorship attribution techniques (Usha et al, 2017). Supervised classification methods are preferred in the current literature (Rocha et al, 2016). In this study, machine learning based approach was used.

Abbasi (2005) collected 20 web forum messages from each 20 authors. Average length was 76.6 words. They used 5 authors and randomly chosen 30 messages in their experiment for comparing feature types and classification techniques. 301 features were chosen and C4.5 and SVM were used. Accuracy for C4.5 was 90% while it was 97% for SVM. Calix et al. (2008) updated an existing C# based stylometry system for verifying author of e-mails. They used 55 style features and K-nearest neighbor algorithm for classification. The average length of e-mails was 150 words.

Layton (2010) evaluated current techniques and identifies some new preprocessing methods. They stated that existing authorship attribution technique SCAP (Source code authorship profile) performs well. A threshold quantity of tweets regarding to attribution task is determined in the paper and informed that 120 tweets per author is an important threshold and there is not a significant improvement in accuracy even in the case of increasing the tweet number greater than the threshold value.

Bhargava, et al (2013) grouped various tweets for increasing the text size under consideration. They prefer to analyze features over a group of tweets instead of a single one. They used syntactic, lexical, tweet specific and emoticon features as author style in which firstly the model was trained by applying SVM as classifier. By increasing the length of each block, they reached 81.42% accuracy for 10 users with 200 tweets each and 77.7% accuracy after increasing tweets number to 250 each. If they increased number of users to 20 with 300 tweets per user, they achieved 64.54% accuracy. Also they informed that while

group of 10 tweets received the best result, using each tweets alone resulted with 7.81% accuracy.

Green, and Sheppard(2013) focused on messages collected from Twitter to analyze most effective feature sets for authorship verification. They used sequential minimal optimization (SMO) algorithm included in Weka for classification 10 authors with 120 tweets from each and had 44% accuracy rates. They compared style makers (SM) feature sets and bag-of-words (BOW) feature sets and informed that SM features are more effective than BOW features for authorship verification. Further, the analysis of the authorship traits for verifying the legitimacy of twitter accounts was examined by Barbon et al (2017). By aiming that, the syntactic, lexical, idiosyncratic and content specific features were applied.

Arakawa et al (2014) investigated a Twitter specific approach which evaluates the category and number of re-tweets. Afroz et al (2014) prepared a large scale study related to posts on forums and malicious search engine optimization. They proposed several features which are suitable to social network messages as word-level bigrams, numbers used in place of letters, capitalization, and existence of foreign words. Azaronyad et al (2015) drew attention to the dynamicity of authors and examined the temporal changes of word usage by authors of tweets and emails and based on this examination they suggested a way to measure the dynamicity of authors' word usage.

Li, et al (2016) used short posts from Facebook. Facebook post, average 20.6 words was applied as the dataset in order to determine whether user is authenticated or not among 30 users in the work. Further, SVM Light with 233 features was applied and 12 tests were conducted. They discussed the challenge of using traditional stylometry on short texts. They examined different feature sets. The success for 10 users with 233 features was 81.6%. When the author number was increased to 20 and 30, the success was slightly dropped to 79.8% and 79.6% respectively (Demir, 2016, 2017).

For the determination of traits in multi authored documents, Macke, and Hirshman (2015) used deep learning techniques that is at the sentence level. The vocabulary and grammatical structure with the application of recurrent neural network model (RNN) is modeled by the authors and it is noticed that application has less performance in the case the number of authors increases.

Schwartz, et al (2013) trained SVM classifier for classification of Twitter messages and n-gram features set was used. The tweets that have fewer than 3 words were removed in the preprocessing process and k-signature of authors that appears in at least k% of author's training set but not appear in others' was defined and used as a feature. Authorship attribution in tweets with a focus on unique signature related with users was studied in the research. In the experiments different number of authors and tweets were used. 65% accuracy was achieved for 50 authors and 500 tweets and 72% was achieved for 1000 tweets.

Decreasing size of submitted data and increasing author number resulted with decreasing the accuracy rate.

Rocha A. et al (2016) compared several algorithms to classify tweets and discussed an extensive review for the existing authorship analysis techniques in micro blogs. They concluded that PMSVM had the best accuracy rate. The success was 48% for 50 authors with 100 tweets. Using more number of tweets increased the accuracy rate; 500 tweets 55% and 1000 tweets 65%. The results offered for the necessity of a plenary method which allows the application of the data context and process it irrespective of its multimodality and further a system which tolerates the lack regarding availability for all author data during training.

Brocardo. et al (2014) proposed a supervised technique used n-gram feature set for authorship identification. They used Enron e-mail dataset. They prepared their data as each block contains 500 characters and each user has 50 blocks. They used 87 users and the EER (equal error rate) was 14.35%. In their late work (2017), they analyzed the use of deep belief networks for authorship verification model of continuous authentication. They achieved 16.73% ERR for 10 users with 140-character-length 100 blocks per user.

An authorship attribution method is offered by Usha et al (2017) in which the tone and personality patterns related with an author is modeled. Method is acquired with the application of convolutional neural network trained on tone and personality data. Data of the authors from Twitter is employed on the models and then psycholinguistic features were united with the final level features. Obtained features were applied for training a linear SVM classifier for prediction of an unknown tweet's author. Their results showed that if data number increased, better results were obtained. However, increasing the number of authors has reverse impact. 15 Users with 250 tweets had 51% accuracy and with 800 tweets results increased to 80% accuracy. However, 50 Users with 250 tweets achieved 50% accuracy and 50 Users with 800 tweets achieved 71% accuracy.

Sirivivasan and Nalini (2017) evaluated the effects of different classification methods for online messages. They used lexical, syntactic, structural and n-gram features and as classifier they examined C4.5, fuzzy classifier and Ada boost classifier. 40 Amazon review messages were collected from each 5 authors and evaluated by using cross validation. Ada boost classifier received the best results with 84% accuracy for 5 authors.

2.A BRIEF NOTE ON ANNS

This brief presentation of artificial neural networks will focus on a particular structure of ANNs, multi-layer feedforward networks, which is the most popular and widely-used network paradigm in many applications including forecasting volatilities and prices in markets. For a general introductory account of ANNs, readers are

referred to Wasserman (1989); Hertz et al. (1991); Smith (1993). Rumelhart et al. (1986a), (1986b), (1994), (1995); Lippmann (1987); Hinton (1992); Hammerstrom (1993); Haykin 1999 illustrate the basic ideas in ANNs.

Hush and Horne (1993) summarize some theoretical developments in ANNs since Lippmann (1987) tutorial article. Masson and Wang (1990) give a detailed description of five different network models. Wilson and Sharda (1992) present a review of applications of ANNs in the business setting. Sharda (1994) provides an application bibliography for researchers in Management Science/Operations Research. A bibliography of neural network business applications research is also given by Wong et al. (1995). Kuan and White (1994) review the ANN models used by economists and econometricians and establish several theoretical frames for ANN learning. Cheng and Titterington (1994) make a detailed analysis and comparison of ANNs paradigms with traditional statistical methods.

Basic structures of artificial neural networks, originally developed to mimic the human brain, are composed of a number of interconnected simple processing elements called neurons or nodes. Each node receives an input signal which is the total ‘‘information’’ from other nodes or external stimuli. The node processes incoming data locally through an activation function and produces a transformed output signal to other nodes or external outputs. Although each individual neuron implements its function rather slowly and imperfectly, collectively a network can perform a surprising number of tasks quite efficiently (Reilly and Cooper, 1990). This information processing characteristic makes ANNs a powerful computational device and able to learn from examples and then to generalize to examples never before seen.

Many different ANN models have been proposed since 1980s. Perhaps the most influential models are the multi-layer perceptrons (MLP), Hopfield networks, and Kohonen’s self organizing networks (Kohonen, 2001). Hopfield (1982) proposes a recurrent neural network which works as an associative memory. An associative memory can recall an example from a partial or distorted version. Hopfield networks are non-layered with complete interconnectivity between nodes. The outputs of the network are not necessarily the functions of the inputs. Rather they are stable states of an iterative process.

2.2. Multi Layer Perceptrons for Clustering

Especially in forecasting the MLP networks are used because of their inherent capability of arbitrary input–output mapping. Other types of ANNs are radial-basis functions networks (Park and Sandberg, 1991, 1993; Chng et al., 1996), ridge polynomial networks (Shin and Ghosh, 1995), and wavelet networks (Zhang and Benveniste, 1992; Delyon et al., 1995) are also very useful in some applications due to their function approximating ability.

An MLP is composed of several layers of nodes. The lowest layer is an input layer where external information is received. The last layer is an output layer where the problem solution is obtained. Hidden layers separate the input layer and output layer. The nodes in adjacent layers are usually fully connected from a lower layer to a higher layer.

2.4 Recurrent Neural Networks (RNN)

In RNNs there are extra neurons in the input layer connected

to the hidden layer and/or output layer just like the other input neurons. These extra neurons hold the contents of one of the layers as it existed when the previous pattern was trained. In this way, the network takes into account previous knowledge it has about previous inputs. These extra neurons are called the context unit and it represents the network’s long-term memory (Balkin 1997).

There are mainly two types of RNNs: Jordan, and Elman recurrent networks. In Jordan recurrent neural network (JNN) additional neurons in the input layer are fed back from output layer (Carcanoa, et al, 2011), while an Elman neural network (ENN) input layer is fed back from hidden layer (Elman, 1990).

2.5 Jordan Recurrent Neural Networks (JNN)

A Jordan neural network (JNN) has several feedback connections from the output layer to the input layer. The input layer has additional neurons, which are fed back from the output layer (Carcanoa, et al, 2011).

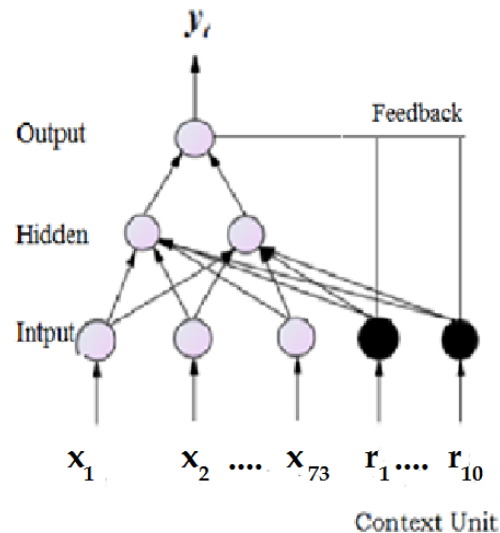


Figure 2. JNN with a single hidden layer representing a nonlinear regression model.

3. DATA

Dataset consists of 17000 tweets collected from Twitter, as 500 tweets for each of 34 authors that meet certain criteria. Raw data collected using the software Nvivo. The

collected raw data is preprocessed in order to obtain same structure and improve classification accuracy. Data preprocessing is a very critical stage for establishing the next stage’s quality. Data in its original form is not in convenient pattern for learning. It needs to be transformed into an appropriate input form. Second step was feature extracting. 200 features in four types are integrated into feature set and used for e-mail authentication. 72 of them are function words which are selected from the list that was prepared by Zheng et al. (2006). The features are extracted by a program in Java, and registered to a text file. Later this text file was reached by our program for training the classifiers and to implement author attribution.

The features that are evaluated are combinations of character-based lexical features, word-based lexical features, syntactic features, structural features and social networking-based features. We collected only textual inputs and did not collect metadata like date of posting, location of user, application for posting, and id. because of the research’s extent. Further, data set is collected without any tendency to any particular content or user.

Studies showed that different types of features have different power of discrimination. Therefore, it is important to identify the key features.

In the second phase we decreased the number of features to 72 by removing the ones which are rare in tweets (Appendix).Nineteen of them are function words in Table 2 below.

Table 2. The nineteen function words

most	most	be	tonight
more	more	thanks	rise
yet	yet	everything	woman
will	will	live	was
every	every	two	

The reason of having that much sparse feature vectors is the nature of tweets which contain few words. Measurements of the features are normalized in the range of 0- 1. Normalization was done by dividing each value by the total word count of the corresponding text, in order to remove the influence of different overall text size.

Feature vectors, created by extracting from Twitter messages, were used as input for modeling artificial neural network (ANN).

4. CLASSIFIERS FOR FIVE TO THIRTY AUTHORS

To train a recurrent artificial neural network that will be able to distinguish tweets of the authors a1, and a2, we choose an appropriate network architecture.

4.1. Network Architecture

The input vector is 72 dimensional, for bias, 1 is added as the first element of each data vector, and we add one component for the recurrent information. Therefore, the neural network will have 74 input neurons. This input vector is multiplied by a 74x74 synaptic weight matrix W1, to create vector of 74 numbers at 74 hidden layer nodes. To this vector 1 is added as the first element as bias. Then this 75 component vector is multiplied by the activation function to involve a nonlinear transformation. Finally another 1x74 synaptic weight matrix W2, multiplies the resulting vector to create a number. Then this number is sent to a “hard limiter” to create +1, or -1 at the output node. This number is also sent as the last component of the input vector to recur at the next iteration.

If the data entered to ANN belongs to a tweet authored by the author a1, and the output is +1, it is OK, otherwise it is erroneous, and synaptic weights must be adjusted by back propagation of the error through iterations, till ANN creates enough correct results at the output node.

4.2. The Case of Five Authors

To classify tweets authored by five authors, 25 pairwise recurrent artificial neural networks are trained, five of which are dummy that are indifferent between +1, and -1. When from a mixture of 500 tweets, 100 from each author are classified by pair wise classifiers, the accuracies of 25 pair wise classifiers are found as in Table 3. The average of pair wise classification accuracies is 89.05%.

Table 3. The accuracies achieved by 25 experts that trained to distinguish tweets by author pairs (ai, aj).

%	a1	a2	a3	a4	a5
a1	0.	92.0	92.5	85.5	90.0
a2	88.0	0.	89.0	84.5	92.0
a3	89.5	89.5	0.	93.0	94.0
a4	86.5	88.0	86.0	0.	86.5
a5	90.5	87.0	89.5	87.5	0.

4. 2.1 Aggregating Expert Votes

An authentication device from these 25 experts for tweets by five authors is created to aggregate their votes. Experts are grouped as competing teams. The team tk, k=1,...,5, of ten experts tk={eik,eki}, i=1,...,5, is trained to distinguish tweets by author ak, from tweets from other four authors. If the data vector v, belonging to a tweet by ak, let k=4, is considered for authentication, most probably the experts e4j, j=1,2,...,5 of the team t4 will rise a flag +1, while other experts ei4, i=1,2,...,5 of the same team will rise flags-1. Since experts of other teams tk, k≠4, are not trained in tweets by a4, they will not be as consistent as the members of the team t4. Therefore, their votes will rather be mixed signals. Upon introduction of a data vector that

belongs to a tweet written by author a4, the votes of 25 experts may be just like the ones in Table 3.

Table 4. The flags raised possibly when a data vector representing a tweet authored by author four entered to the authentication system.

no	1	2	3	4	5
1	0	-1	1	1	-1
2	-1	0	-1	-1	1
3	1	-1	0	-1	-1
4	-1	1	1	0	1
5	1	1	1	-1	0

Let us reorganize the votes of five competing teams as in Table 4. In rows, +1s, in columns -1s are correct votes. The team that has more correct votes, wins the competition. In this case the highest number of true votes, (6) is collected by Team 4. Hence we conclude that the tweet whose data vector is entered, is authored by a4.

Table 5. Team 4, has the largest number of correct votes, and wins the competition.

T	V	Rows					Columns				
1	4	0	-1	1	1	-1	0	-1	1	-1	1
3	2	1	-1	0	-1	-1	1	-1	0	1	1
4	6	-1	1	1	0	1	1	-1	-1	0	-1
5	5	1	1	1	-1	0	-1	1	-1	1	0

Aggregating pairwise classification votes as in the above, 500 shuffled tweets of five authors are classified. The percentage of true positives are as in Table 6, where the average accuracy is 82%.

Table 6. The accuracies in distinguishing shuffled tweets authored by five authors. The average is 82.2%

	a1	a2	a3	a4	a5	M
%	74	78	86	87	86	82.2

4. 3. The Case of Ten Authors

To classify tweets authored by ten authors, 100 recurrent artificial neural networks are trained, ten of which are dummy that are indifferent between votes +1, and -1. When from a mixture of 1000 tweets, 100 from each author are classified by pairwise classifiers, the accuracies of 100 pairwise classifiers are found as in Table 7.

Table 7. The accuracies achieved by 100 experts that trained to distinguish tweets by author pairs (ai, aj).

%	a1	a2	a3	a4	a5	a6	a7	a8	a9	a10
a1	0	84	86	77	96	89	95	89	78	80
a2	88	0	88	82	88	90	92	93	92	72
a3	86	79	0	78	88	90	95	90	96	78
a4	74	83	82	0	80	86	88	91	92	67
a5	84	82	90	85	0	80	88	92	86	78
a6	83	87	85	82	73	0	84	84	86	72
a7	92	92	86	94	90	91	0	93	92	92
a8	82	91	89	89	80	82	93	0	88	86
a9	74	92	93	80	86	91	89	92	0	86
a10	86	81	78	73	58	71	83	78	72	0

The average pair wise authentication accuracy is 88.99%

4.3. 1 Aggregating Expert Votes

To create a authentication device from these 100 experts for tweets by 10 authors, authentication is made on a single tweet as in the case of five authors. We call this technique of aggregating decisions as competing artificial neural network teams (CANNT). In Table 8, the accuracy of 100 artificial neural networks to distinguish shuffled tweets by 10 authors is given.

Table 8. The accuracies in distinguishing shuffled tweets authored by 10 authors.

	a1	a2	a3	a4	a5
%	73	77	84	66	76
	a6	a7	a8	a9	a10
%	76	88	89	79	83

The average authentication accuracy of a single tweet is 79.1%.

4. 4. The Case of Fifteen Authors

To classify tweets authored by fifteen authors, 225 recurrent artificial neural networks are trained, fifteen of which are dummy that are indifferent between votes +1, and -1. When from a mixture of 1500 tweets, 100 from each author are classified by pairwise classifiers, the average accuracy of 225 pairwise classifiers is 88.13%.

4. 4.1 Aggregating Expert Votes

To create a authentication device from these 225 experts for tweets by 15 authors, authentication is made on a single tweet as in the case of five authors. We call this technique of aggregating decisions as competing artificial neural network teams (CANNT). In Table 9, the accuracy of 225 artificial neural networks to distinguish shuffled tweets by 15 authors is given.

Table 9. The accuracies in distinguishing shuffled tweets authored by 15 authors.

	a1	a2	a3	a4	a5
%	66	76	81	77	74
	a6	a7	a8	a9	a10
%	72	91	86	85	67
	a11	a12	a13	a14	a15
%	68	79	81	65	75

The average authentication accuracy of a single tweet is 76.2%.

4. 5. The Case of Twenty-Twenty Five-Thirty-Thirty-Four Authors

To classify tweets authored by twenty-twenty five-Thirty-thirty four authors, for N is the number of authors, N×N recurrent artificial neural networks are trained, N of which are dummy that are indifferent between votes +1, and -1.

When from a mixture of $N \times 100$ tweets are classified by pairwise classifiers, the accuracies of $N \times N$ pairwise classifiers are found as in Table 10. The average authentication accuracies of a single tweet are also shown in the same table.

Table 10. The pair wise accuracies and accuracies in distinguishing single tweets authored by N authors.

N	5	10	15	20	25	30	34
Pair wise	89	89	88	88	88	88	88
Single	82	79	76	74	71	68	69

4. 5.1 Aggregating Expert Votes

To create a authentication device from these $N \times N$ experts for tweets by N authors, authentication is made on a single tweet. In Tables 11-14, the accuracies of $N \times N$ artificial neural networks to distinguish shuffled tweets by N authors are given.

Table 11. The accuracies in distinguishing shuffled tweets authored by 20 authors.

	a1	a2	a3	a4	a5
%	66	66	75	65	70
	a6	a7	a8	a9	a10
%	77	86	88	82	50
	a11	a12	a13	a14	a15
%	63	75	71	68	80
	a16	a17	a18	a19	A20
%	64	77	81	72	69

The average authentication accuracy of a single tweet is 74%.

Table 12. The accuracies in distinguishing shuffled tweets authored by 25 authors.

	a1	a2	a3	a4	a5
%	65	72	77	64	67
	a6	a7	a8	a9	a10
%	69	66	87	80	58
	a11	a12	a13	a14	a15
%	67	76	70	56	72
	a16	a17	a18	a19	A20
%	68	80	80	82	63
	a21	a22	a23	a24	a25
%	72	67	74	68	79

The average authentication accuracy of a single tweet is 71%.

Table 13. The accuracies in distinguishing shuffled tweets authored by 30 authors.

	a1	a2	a3	a4	a5
%	69	72	80	53	75
	a6	a7	a8	a9	a10
%	56	61	81	79	75
	a11	a12	a13	a14	a15
%	54	80	63	53	80
	a16	a17	a18	a19	A20
%	67	72	70	70	67
	a21	a22	a23	a24	a25
%	62	65	61	66	77
	a26	a27	a28	a29	a30
%	69	63	65	63	68

The average authentication accuracy of a single tweet is 68%.

Table 14. The accuracies in distinguishing shuffled tweets authored by 34 authors.

	a1	a2	a3	a4	a5
%	66	70	70	68	70
	a6	a7	a8	a9	a10
%	70	73	82	69	55
	a11	a12	a13	a14	a15
%	51	73	68	48	71
	a16	a17	a18	a19	A20
%	63	74	66	79	62
	a21	a22	a23	a24	a25
%	66	57	64	68	80
	a26	a27	a28	a29	a30
%	76	64	64	61	67
	a31	a32	a33	a34	
%	80	74	67	64	

The average authentication accuracy of a single tweet is 67%.

CONCLUSION

Dataset consists of 17000 tweets collected from Twitter, as 500 tweets for each of 34 authors that meet certain criteria. Raw data is collected by using the software Nvivo. The collected raw data is preprocessed to extract frequencies of 72 features. When N authors are selected, since artificial neural networks are more successful distinguishing two classes, $N \times N$ neural networks are trained for pairwise classification. These experts then organized as a special team (CANNT) to aggregate decisions of these N competing teams of experts. For several numbers of authors, authentication accuracies are shown in Table 9.

Classification accuracies are shown in Tables 5-8, and 10-13. Using committee machines for classification jobs, these accuracies may be improved around 10% in average.

REFERENCES

- Abbasi, A. and Chen, H. (2005) "Applying Authorship Analysis to Extremist-Group Web Forum Messages," the IEEE Computer Society, pp. 1541-1672.
- Abbasi, A. and Chen, H. (2008), Writeprints: A Stylometric Approach to Identity-Level Identification and Similarity Detection in Cyberspace, ACM Transactions on Information Systems, vol. 26, no. 2, pp7.
- Afroz,S., Brennan,M. and Greenstadt R. (2012), Detecting hoaxes, frauds, and deception in writing style online. In Proceedings of the 33rd conference on IEEE Symposium on Security and Privacy. IEEE.
- Arakawa, Y., Kameda, A., Aizawa, A., Suzuki, T. (2014), Adding Twitter-Specific Features to Stylistic Features for Classifying Tweets by User Type and Number of Retweets, Journal of The Association For Information Science and Technology, 65(7):1416–1423.
- Azarbonyad, H., Dehghani, M. and Marx, M. (2015) Time-Aware Authorship Attribution for Short Text Streams, Proceedings of the 38th International ACM SIGIR Conference on Research and Development in Information Retrieval, pp727-730.
- Balkin, S.D. (1997). Using recurrent neural networks for time series forecasting, Working Paper Series number 97–11, International Symposium on Forecasting, Barbados 2–54.
- Barbon, Jr, S., Igawa, R.A., Zarpeleo, B.B., (2017) Authorship verification applied to detection of compromised accounts on online social networks, Multimedia Tools and Applications, vol. 76, no. 3, pp 3213-3233
- Bhargava, M., Mehndiratta, P, and Asawa , K., (2013) Stylometric Analysis for Authorship Attribution on Twitter, Big Data Analytics, pp 37-47.
- Brocardo, M. L., Traore, I. Saad, S., Woungang, I., (2013) Authorship Verification for Short Messages Using Stylometry, in Proc. of the IEEE Intl. Conference on Computer, Information and Telecommunication Systems (CITS), pp 1-6.
- Brocardo, M. L., Traore, I. Saad, S., Woungang, I., (2014) Verifying Online User Identity using Stylometric Analysis for Short Messages, Journal of Networks, Vol. 9, No. 12, pp 3347-3355.
- Brocardo, M.L, Traorea, I., Woungang, I., (2015), Authorship Verification of E-Mail and Tweet Messages Applied for Continuous Authentication, Journal of Computer and System Sciences, vol. 81, no. 8, pp. 1429-1440.
- Brocardo, M. L., Traore, I., Woungang, I., Obaidat, M. S., (2017) Authorship Verification Using Deep Belief Network Systems, International Journal of Communication Systems vol. 30, no.12, e3259.
- Brocardo, C.M., and White, H. (2007). Artificial neural networks: An econometric perspective, Econometric Reviews, 13, 1–92
- Brocardo, M. L., Traore, I., Woungang, I., Obaidat, M. S., (2017) Authorship verification using deep belief network systems, International Journal of Communication Systems 30(12):e3259
- Can, M. (2012) Teaching Neural Networks to Detect the Authors of Texts Using Lexical Descriptors, Southeast Europe Journal of Soft Computing, vol. 1, no. 1, pp 57-72.
- Carcanoa, E.C., Bartolinia, P., Muselli, ., and Piroddi, L., (2008) Jordan recurrent neural network versus IHACRES in modelling daily streamflows, Journal of Hydrology, Volume 362, Issues 3–4, 5 December 2008, Pages 291-307
- Cheng, ., and Titterington, D. M, (1994) Neural Networks: A Review from a Statistical Perspective, Statistical Science, Volume 9, Number 1 2-30.
- Chng, E.S., Chen, S., Mulgrew, B. (1996) Gradient radial basis function networks for nonlinear and nonstationary time series prediction. IEEE Transactions on Neural Networks 7 (1), 190–194.
- Delyon, B., Juditsky, A., Benveniste, A. (1995) Accuracy analysis for wavelet approximations. IEEE Transactions on Neural Networks 6 (2), 332–348.
- Demir, N. M. (2016), Authorship Authentication for Twitter Messages Using Support Vector Machine, Southeast Europe Journal of Soft Computing, Vol. 5, No. 2, pp. 1-6.
- Demir, N.M. (2017), Better Features Sets for Authorship Attribution of Short Messages, Southeast Europe Journal of Soft Computing, Vol. 6, No. 1, pp. 8-12.
- Elman, J. L. (1990), 'Finding structure in time', Cognitive Science, 14, 179-211.
- Green R, Sheppard J. (2013), Comparing Frequency and Style-Based Features for Twitter Author Identification, Proc. Twenty-Sixth International Florida Artificial Intelligence Research Society Conference, pp. 64-69.
- Hamilton, A., Madison, J., and Jay,J., Goldman, L., (2008) The Federalist Papers, Oxford World's Classics.
- Hammerstrom, D. (1993) Neural networks at work, IEEE Spectrum, June, 26–32.
- Haykin, S. (1999) Neural Networks, a Comprehensive Foundation, Prentice Hall.
- Hertz, J., Krogh, A., Palmer, R.G., (1991) Introduction to the Theory of Neural Computation. Addison-Wesley, Reading, MA.
- Hinton, G. E., (1992) How Neural Networks Learn from Experience, Scientific American, September 1992, pp. 145-151.
- Hopfield, J.J. (1982) Neural networks and physical systems with emergent collective computational abilities. Proceedings of the National Academy of the Sciences of the U.S.A. 79, 2554– 2558.

- Hush, D.R., Horne, B.G. (1993) Progress in supervised neural networks: What's new since Lippmann? *IEEE Signal Processing Magazine*, January, 8–38.
- Kohonen, T. (2001) *Self-Organizing Maps* 3rd Edition, Springer, 2001.
- Kuan, C.-M., Liu, T. (1995) Forecasting exchange rates using feedforward and recurrent neural networks. *Journal of Applied Econometrics* 10, 347–364.
- Kuan, C.M., Hornik, K. and H. White (1994) A convergence result for learning in recurrent neural networks', *Neural Computation*, 6, 620-640.
- Kuan, C.M. and H. White (1994) Adaptive learning with nonlinear dynamics driven by dependent processes, *Econometrica*, vol.62 No 5 pp. 1087-1114.
- Layton, R., Watters, P. and Dazeley, R. (2010) Authorship Attribution for Twitter in 140 Characters or Less, *Second Cybercrime and Trustworthy Computing Workshop*, pp. 1-8.
- Lippmann, R.P., (1987) An introduction to computing with neural nets, *IEEE ASSP Magazine*, April, 4–22.
- Macke, S., and Hirshman J. (2015) *Deep Sentence-Level Authorship Attribution*, Stanford University.
- Masson, E., Wang, Y.-J., (1990) Introduction to computation and learning in artificial neural networks. *European Journal of Operational Research* 47, 1–28.
- Park, J., Sandberg, I.W., (1991) Universal approximation using radial basis function networks. *Neural Computation* 3, 246– 257.
- Park, J., Sandberg, I.W. (1993) approximation and radial basis function networks. *Neural Computation* 5, 305–316.
- Reilly, D.L., Cooper, L.N., (1990) An overview of neural networks: early models to real world systems. In: Zornetzer, S.F., Davis, J.L., Lau, C. (Eds.), *An Introduction to Neural and Electronic Networks*. Academic Press, New York, pp. 227–248.
- Rocha, A. et al. (2016) Authorship Attribution for Social Media Forensics, in *IEEE Transactions on Information Forensics and Security*, vol. 12, no. 1, pp. 5-33
- Rumelhart, D.E., Hinton, G.E., Williams, R.J. (1986a) Learning representations by backpropagating errors. *Nature* 323 (6188), 533–536.
- Rumelhart, D.E., Hinton, G.E., Williams, R.J. (1986b) Learning internal representation by back-propagating errors. In: Rumelhart, D.E., McClelland, J.L., the PDP Research Group (Eds.), *Parallel Distributed Processing: Explorations in the Microstructure of Cognition*. MIT Press, MA.
- Rumelhart, D.E., Widrow, B., Lehr, M.A. (1994) The basic ideas in neural networks. *Communications of the ACM* 37 (3), 87–92.
- Rumelhart, D.E., Durbin, R., Golden, R., Chauvin, Y. (1995) Backpropagation: the basic theory. In: Chauvin, Y., Rumelhart, D.E. (Eds.), *Backpropagation: Theory, Architectures, and Applications*. Lawrence Erlbaum Associates, New Jersey, pp. 1–34.
- Schwartz, R. et al. (2013) Authorship Attribution of Micro-Messages, *Proceedings of the 2013 Conference on Empirical Methods in Natural Language Processing*, pp 1880-1891.
- Sharda, R., (1994) Neural networks for the MS/OR analyst: An application bibliography. *Interfaces* 24 (2), 116–130.
- Sharda, R., Patil, R.B. (1992) Connectionist approach to time series prediction: An empirical test. *Journal of Intelligent Manufacturing* 3, 317–323.
- Shin, Y., Ghosh, J. (1995) Ridge polynomial networks. *IEEE Transactions on Neural Networks* 6 (3), 610–622.
- Smith, M. (1993) *Neural Networks for Statistical Modeling* Thomson Learning ©1993
- Srinivasan, L. & Nalini, C. (2017) An improved framework for authorship identification in online messages, *Cluster Comput.* <https://doi.org/10.1007/s10586-017-1563-3>
- Usha, A. and Thampi, S. (2017) Authorship Analysis of Social Media Contents Using Tone and Personality Features, In: Wang G., Atiquzzaman M., Yan Z., Choo KK. (eds) *Security, Privacy, and Anonymity in Computation, Communication, and Storage. SpaCCS 2017. Lecture Notes in Computer Science*, vol 10656.
- Wasserman, P.D. (1989) *Neural Computing: Theory and Practice*. Van Nostrand, Reinhold, New York.
- Wilson, R., Sharda, R. (1992) Neural networks. *OR/MS Today*, August, 36–42.
- Wong, S.Q., Long, J.A. (1995) A neural network approach to stock market holding period returns. *American Business Review* 13 (2), 61–64.
- Zhang, Q., Benveniste, A. (1992) Wavelet networks. *IEEE Trans-actions on Neural Networks* 3 (6), 889–898.
- Zheng, R. Li, J., Chen, H. and Huang, Z. (2006), “A Framework for Authorship Identification of Online Messages: Writing-Style Features and Classification Techniques”, *Journal of the American Society for Information Science and Technology*, vol. 57, no. 3, pp. 378–393

PPENDIX: 72 Features

Reference Number	Features	Description
1	Total number of characters	
2	Total number of characters with space	
3	Total number of words	
4	Average word length	
5	Total number of sentences	
6	Total number of uppercase	
7	Total number of lowercase	
8	Total number of short words	Less than four characters
9-20	Frequency of different length of words	(ie.3 characters)
21-34	Frequency of letters	A-Z except; bcdehmpux
35-36	Frequency of emotions	ie.Smiley face”:)”;sad face :(
37	Frequency of ‘LOL’	
38	Total number of special characters	@%\$&*#
39	Total number of punctuations	.!/?;
40	Frequency of Vowels	A,e,o,u,i
41	Frequency of GDeterminers	ie. a, an, any,another, other
42	Frequency of SDeterminers	ie. my,mine, his, her
43	Frequency of Interjections	ie. aah, hmm, jeez, oops
44	Frequency of ObjectPronoun	ie. hers, us, our, yours
45	Frequency of SubjectPronoun	ie. I, you, he, she, it
46	Frequency of IntensivePronoun	ie. myself, yourself, herself
47	Frequency of To be	am, is, are, ‘m, ‘s, ‘re
48-51	Frequency of Some characters	+ - , ...
52	Frequency of digits	ie. 1,2,3
53	Frequency of Brackets	ie. {, }, [
54-72	Frequency of function words	19 words