Sentiment Analysis of Serious Suicide References in Twitter Social Network

Wael Korani[a] and Malek Mouhoub[b]

Department of Computer Science, University of Regina, Regina, Saskatchewan, Canada

Keywords: Sentiment Analysis, Twitter, Suicide Thoughts, Artificial Intelligence.

Abstract: Sentiment analysis analyzes people emotions, attitudes, and opinion towards organizations, services, issues, and individuals. Opinions are the core of almost all human activities because they consider a significant influencers of our behaviors. With the growing popularity of social media applications (micro-blogs, twitter, comments, etc), users of these platforms express their emotions through their posts and comments. Suicide is one of these dangerous emotions that threaten the public health of Canadians, and mortality form suicide is the third leading cause of death in teenage. In this paper, we propose a suicide classifier system called Auto Twitter Suicide Detector System (ATSDS) that provides support to authorities to take appropriate actions in order to protect communities from such kind of thoughts. The proposed twitter suicide detector system is a classifier system using data gathered from twitter to detect those related to suicide. Our system is built using deep neural network on multi-purpose cluster computing system called spark. In order to asses the system performance, in terms of accuracy, we have conducted several experiments and tuned neural network parameters to achieve higher performance. The results returned are very promising.

1 INTRODUCTION

The International Statistical Classification of Diseases and Related Health Problems (ISCDRHP) refers to suicide related behavior as ”intentional self-harm”. Suicide-related behavior includes thoughts, behaviors, and communications related to suicide. In (Yip et al., 2003), Yip et al. defined suicide related ideation as thoughts of ending one’s life or a wish to be dead. A suicide attempt is one form of self-injury whereby the attempt is to end one’s life. There are two types of suicides: active and passive suicide. Active suicide is an effective way of suicide and gives a slim chance of interruption, such as hanging, shooting, and jumping (Glass Jr and Reed, 1993). However, passive suicide is a less violent way of suicide that allows intervention, such as overdose, poisoning, and international malnutrition, which is called indirect self destructive behaviors (ISDBs). In (Conwell et al., 1996), Conwell defined ISDBs as ”an act of omission or commission that causes self-harm leading indirectly, over time, to the patient’s death”. ISDBs are common among older adults who have suicide signs, such as refusing to eat or drink and failing to take medications (Brown et al., 2004).

Canadian Vital Statistics Death (CVSD) is responsible for reporting the cause of death in Canada, which is an effective mechanism for monitoring the death by suicide in Canada. In 2005, Public Health Agency of Canadian Suicide reported on their website that suicide was the eighth leading cause of death for adults between (55-64) years (13.0 per 100,000). In (Buchanan et al., 2006), Canadian Coalition for Seniors Mental Health reported that older adults have the highest rate of death by suicide across all age groups. In (Navaneelan, 2012), Navaneelan reported that the suicide rate in Canada declined from 12.7 per 100,000 between 1989 and 1992, down to 11.5 per 100,000 in 2009.

In the last couple of decade, social media plays a crucial role in our social life. Most people want to be in groups, where they can share ideas, experiences, emotions, etc. Social media applications help people share their ideas, problems, get solutions from other like-minded people. In (Kaplan and Haenlein, 2010), Kaplan and Haenlein defined social media as ”a group of internet-based applications that are built on the ideological and technological foundation of Web 2.0 and that allow the creating and exchange of user gener-

[a] https://orcid.org/0000-0002-1419-1149
[b] https://orcid.org/0000-0001-7381-1064

Korani, W. and Mouhoub, M.
Sentiment Analysis of Serious Suicide References in Twitter Social Network.
DOI: 10.5220/0008894003390346
In *Proceedings of the 12th International Conference on Agents and Artificial Intelligence (ICAART 2020) - Volume 2*, pages 339-346
ISBN: 978-989-758-395-7; ISSN: 2184-433X

ated content". Users interact and share negative and positive experience and learn from each other though social media. Social media is available for any user at any time. There is no limit in time and space on social media, and users can share information at any time and spread it in a second.

In 2006, Twitter was developed in a different way of Facebook (Carlson, 2011). Twitter is another popular, widespread, and limited social network. It is a micro-blogging that gives only 140 characters for each message. An instance twitter message is called tweet, and twitter friends are called followers. Posted tweets from users' friends will be shown on user's profile page. Users on Facebook and Twitter can post text, photo, link, or video. Twitter gives a user the ability to create an instance message that introduces an idea without any barriers. As the third quarter of 2016, the number of active Twitter users was growing each month, which was estimated to be around 317 million active users each month. Twitter becomes more popular and has around 500 million instance message every day.

Sentiment analysis can be performed using different machine learning approaches. Pang suggested that the current research on sentiment analysis focuses on two major things: to identify the given text whether it is subjective or objective. In addition, it may identify the polarity of the subjective texts (Pang et al., 2008). Sentiment analysis has been used for range of topics, such as movie review, products or services reviews, political opinion, and emotions. In this paper, we focus on sentiment analysis related to suicide thoughts. Sentiment analysis was conducted on suicide thoughts that have been reported using written communication of suicide on the Web via bulletin boards (Ikunaga et al., 2013). In (Matykiewicz et al., 2009), unsupervised machine learning was also implemented to distinguish between actual suicide notes and newsgroups. Suicide thoughts are also released in chat rooms with no restrictions (Becker and Schmidt, 2005).

Social media, specially twitter, along with sentiment analysis play a significant role in improving the suicide research by analyzing individuals activities through their posts. In this paper, twitter social media is used to build a classifier system, which is called Auto Twitter Suicide Detector System (ATSDS). The proposed system, TSDS, is capable of detecting twitter users who have suicide thoughts or interested in the suicide topic. The proposed system is built on multi-purpose cluster computing system called spark along with deep neural network. The accuracy of different models are evaluated to choose the best parameter for the neural network.

2 RELATED WORK

Few studies were conducted to use classification approaches to automatically identify suicide-related communications in twitter and other social media. Studies showed a strong positive correlation between suicide rates and the volume of social media posts and comments that related to suicide thoughts (Won et al., 2013). In (Won et al., 2013), Won et al. concluded that the social media data may help in national suicide forecasting and preventing. Jashinsky suggested that there is a relationship between suicide risk and twitter conversation (Jashinsky et al., 2014). John et al. analyzed twitter posts the 24 hours prior to the death by suicide (Gunn and Lester, 2015). The results showed that persons who committed suicide have positive emotions over the last 24 hours and a change in focus from the self to others. Although the study conducted over one case study, the authors later on used more cases. The authors used the Linguistic Inquiry and Word Count (LIWC) software to identify emotional words (Pennebaker et al., 2001).

In (Poulin et al., 2014), Poulin et al. conducted an experiment on a group of US war veterans who shared their Twitter and Facebook over time. The authors proposed a suicide prediction system based on clinical notes of US war veterans, and the system showed high performance (60% accuracy). In addition, the authors concluded that persons who recorded fear, agitation, and delusion behaviors had committed suicide. In (Sueki, 2015), Sueki conduced an experiment using posts of Twitter users to find the relationship between suicide-related tweets and suicidal behavior. The results showed that some particular phrases such as "want to commit suicide" was strongly associated with lifetime suicide attempts. However, phrases that suggest suicide intent, such as "want to die" have less strong association with suicide, because such phrases could be used when a person had a bad day. In (Abboute et al., 2014), Abboute et al. proposed a system to classify "risky" and "non risky" tweets with accuracy 60%. The authors concluded a number of emotions related to suicide, such as hurt, bulling, and insults in the "risky" category.

3 DATASET OVERVIEW

Although a few studies were conducted to predict suicide thoughts using machine learning models, there is no reliable dataset was publicly published. Reliable dataset is one of the challenges in creating a machine learning model. The main reason that there is no existing reliable dataset is that there is no agreement about specific features to characterize suicide notes. Thus, our dataset is pulled from Twitter and distributed in files. The dataset has 1719 files that contain 815871 tweets from different regions specially Canada. Tweets in our dataset are raw data that we should firstly clean.

Our dataset has attributes, such as username, country, time, location, posted message, etc. Tweets, country, and city are all attributes that we need to create our system. Tweets are filtered based on the content of "suicide" word, and the resulted tweets are divided into two classes "suicide" and "non suicide". The results show that the dataset includes 368 tweets in suicide class. We then extended the filter using some extra words that might have relationship with the suicide thoughts in literature, such as killing myself, hate myself, hate this life, want to die, and hate people. The results show more tweets belongs to suicide class, which has 469 tweets. Finally, we mix the suicide tweets with a new class of non suicide class that has 1407 tweets. The entire dataset includes 1876 tweets. Figure 1 shows the histogram of the suicide and non suicide classes. Figure 2 and 3 are two samples tweets of each class of these classes: suicide and non suicide.

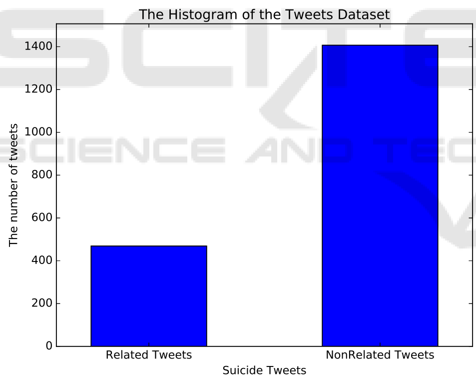
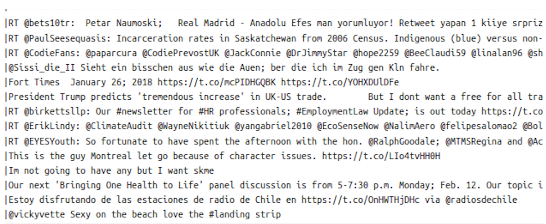


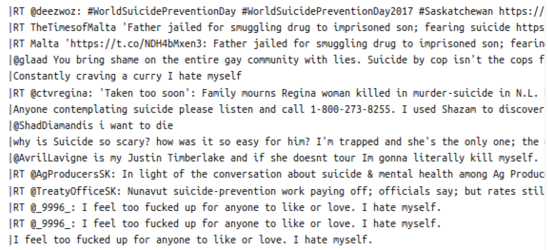
Figure 1: Histogram of suicide and non suicide tweets.



|RT @bets10tr: Petar Naumoski; Real Madrid - Anadolu Efes man yorumluyor! Retweet yapan 1 kiiye srpriz
|RT @PaulSeesequasis: Incarceration rates in Saskatchewan from 2006 Census. Indigenous (blue) versus non-
|RT @CodieFans: @paparcura @CodiePrevostUK @JackConnie @DrJimmyStar @hope2259 @BeeClaudi59 @linalan96 @sh
|@Sissi_die_II Sieht ein bisschen aus wie die Auen; ber die ich im Zug gen Kln fahre.
|Fort Times January 26; 2018 https://t.co/mcPIDHGQBK https://t.co/YOHXDUlDFe
|President Trump predicts 'tremendous increase' in UK-US trade. But I dont want a free for all tra
|RT @birkettsllp: Our #newsletter for #HR professionals; #EmploymentLaw Update; is out today https://t.co
|RT @ErikLindy: @ClimateAudit @WayneNikitiuk @yangabriel2010 @EcoSenseNow @NalimAero @felipesalomao2 @Bol
|RT @EYESYouth: So fortunate to have spent the afternoon with the hon. @RalphGoodale; @MTMSRegina and @Ac
|This is the guy Montreal let go because of character issues. https://t.co/LIo4tvHH0H
|Im not going to have any but I want skme
|Our next 'Bringing One Health to Life' panel discussion is from 5-7:30 p.m. Monday; Feb. 12. Our topic i
|Estoy disfrutando de las estaciones de radio de Chile en https://t.co/OnHWTHjDHc via @radiosdechile
|@vickyvette Sexy on the beach love the #landing strip

Figure 2: Sample of non suicide tweets.

Figure 4 and 5 show top regions in our dataset. Figure 4 shows the top seven regions in the suicide class. It shows that Regina/Canada, Saskatoon/Canada, and Saskatchewan/Canada are the top three regions in our dataset. However, Figure 5 shows the top eight regions in the non suicide class. Regina/Canada, Saskatoon/Canada, and Saskatchewan/Canada represent big share in our dataset. In addition, the dataset contains significant number of tweets from Malta and Philadelphia/USA in both classes.



|RT @deezwoz: #WorldSuicidePreventionDay #WorldSuicidePreventionDay2017 #Saskatchewan https://
|RT TheTimesofMalta 'Father jailed for smuggling drug to imprisoned son; fearing suicide https
|RT Malta 'https://t.co/NDH4bMxen3: Father jailed for smuggling drug to imprisoned son; fearin
|@glaad You bring shame on the entire gay community with lies. Suicide by cop isn't the cops f
|Constantly craving a curry I hate myself
|RT @ctvregina: 'Taken too soon': Family mourns Regina woman killed in murder-suicide in N.L.
|Anyone contemplating suicide please listen and call 1-800-273-8255. I used Shazam to discover
|@ShadDiamandis i want to die
|why is Suicide so scary? how was it so easy for him? I'm trapped and she's the only one; the
|@AvrilLavigne is my Justin Timberlake and if she doesnt tour Im gonna literally kill myself.
|RT @AgProducersSK: In light of the conversation about suicide & mental health among Ag Produc
|RT @TreatyOfficeSK: Nunavut suicide-prevention work paying off; officials say; but rates stil
|RT @_9996_: I feel too fucked up for anyone to like or love. I hate myself.
|RT @_9996_: I feel too fucked up for anyone to like or love. I hate myself.
|I feel too fucked up for anyone to like or love. I hate myself.

Figure 3: Sample of suicide tweets.

The entire dataset will be used in training and testing processes of creating the model, which causes a bias in the model. We expect to get high accuracy using this method. Then, the dataset will be divided into two parts: training and testing dataset. The training dataset represents 75% of the total dataset. The training dataset will be used to create the model. The rest of the dataset that represents 25% will be used in testing the accuracy of the model. The testing dataset is used to evaluate the performance of the model. The accuracy of the neural network model is used to evaluate the performance of our proposed model.

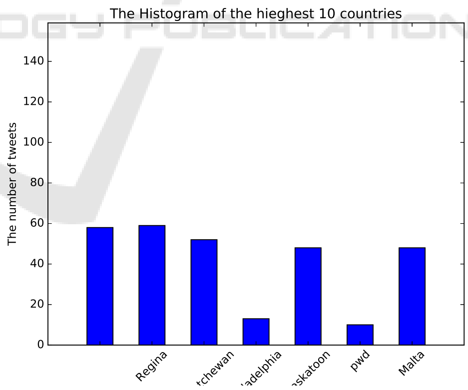


Figure 4: Top seven regions of suicide class.

Second stage, 10-fold cross validation approach is used, which is a recommended technique to avoid producing a bias model. In this stage, the dataset is divided into 10 parts where nine parts is used in training a model and one part is used in testing the created model. This process is repeated for all combinations of train-test splits. Figure 6 shows the process of five folds cross validation. The cross validation process is better than dividing the dataset into specific training and testing data as shown in first stage.

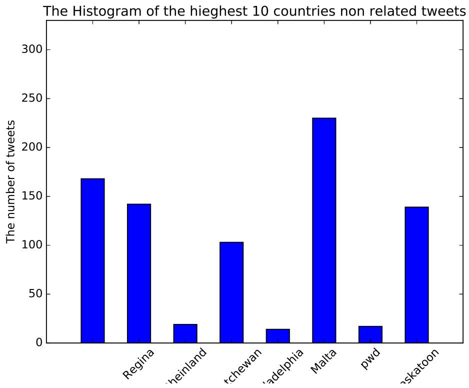


Figure 5: Top ten regions for the non related suicide tweets.

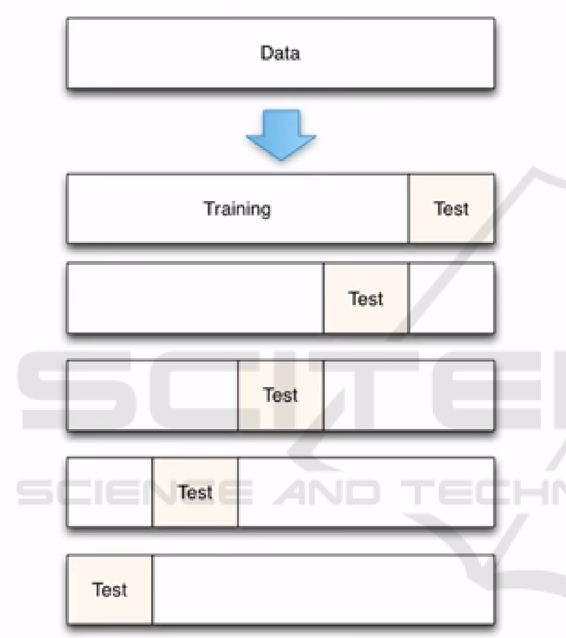


Figure 6: 5-fold Cross Validation Example.

4 PROPOSED MODEL

Our model is created using deep forward neural network on Spark platform using Scala language as shown in Figure 11. Spark is presented in this paper for sentimental analysis of Twitter suicide posts. It is an open source engine multi-purpose cluster computing system for data processing. Spark is used in many applications and among them machine learning applications. It has MLlip library that provides a machine learning functionality, such as classification, clustering, regression, and prediction. MLlib has two packages mllip (built on the top fo RDD) and ml (built on the top of the dataframes). Spark is used for all the operations that were implemented in this paper such as training, cross validation, pipelines, classifying, and computing classifier performance. These operations reveal better understanding of the created model. Parameters of the created model should be tuned to find the best values that improve the accuracy of the model.

4.1 Spark Core

Spark is an open source cluster computing developed by UC Berkely AMPLap. In 2010, Spark is adopted by Apache Software Foundation. Apache Spark is an open source engine multi-purpose cluster computing system for data processing on a large scale. It provides fast memory computing, and it consists of high level tools such as Spark streaming, data frames, SQL, MLlip for machine learning and GraphX for graph processing as shown in Figure 7. The core engine of Spark provides monitoring, scheduling, and distributing of application across the computing cluster. Spark is implemented in Scala language, which runs on (JVM) Java Virtual Machine.

Spark has some great features: Spark API is available in different languages, such as Scala, Java, Python, and R. It runs on a web user interface for checking, monitoring, results, and Spark jobs (Karau et al., 2015). In the last few years, Spark becomes very popular among the companies, such as eBay, Yahoo, Amazon, Databrickes, Baidu, TripAdvisor, and others.

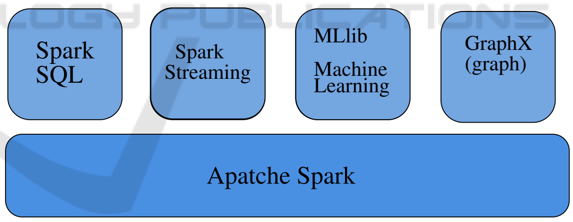


Figure 7: Spark Stack.

4.2 Artificial Neural Network

In 1943, McCulloch proposed the first mathematical model of a neuron (McCulloch and Pitts, 1943). In 1958, Rosenblatt proposed the first neural network known as perceptron (Rosenblatt, 1958). In 2002, Yu Hen proposed a mathematical computing paradigm called artificial neural network that models the operations of biological neural system (Hwang and Hu, 2001). The building block of any neural network model is the neuron. The neuron model that proposed by McCulloch is the most widely used neuron, and the multilayer perceptron is the most widely neural network, which consists of several sequential connected layers of perceptrons.

There are several types of neural network, such as feed-forward and recurrent networks. In the feed-forward networks, the output signal of a neuron has no influence on its inputs. However, the recurrent networks, the output signals of neurons are feedback given as their input signals. The multilayer perceptron that has been used in our project is the feed-forward networks.

4.2.1 Neuron Model

A neuron consists of net function and activation function (transfer function). Figure 9 shows few activation functions that have been used in literature. However, the net function is used to determine how the input signals are combined inside the neuron. The formula for net function is:

$$u = \sum_{i=0}^{N} x_i w_i \tag{1}$$

where w is the weight, and w_0 is the threshold and its corresponding input x_0 is always equal one. In addition, the input x_0 does not form a connection between two neurons as others do. The output of neuron is denoted by Y, which is the output of the net function u by one of the activation function list in Figure 9.

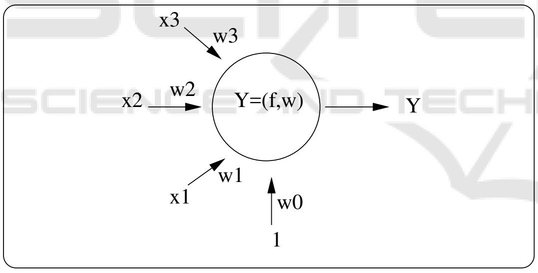


Figure 8: Neuron Model.

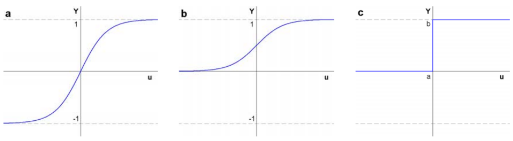


Figure 9: Commonly used transfer functions a - hyperbolic tangent , b - logistic sigmoid , c - threshold.

4.2.2 Multilayer Perceptron Model

A single layer perceptron is able to classify only linearly separable data. A multilayer perceptron (MLP) is a network that includes two or three layers of neurons as shown in Figure 10. MLP consists of one input layer and one output layer, and one or more hidden layers. The MLP network is considered a fully connected if every node in a given layer is connected to every node in the next layer. It is used in many applications, because it has the ability to solve problems that do not have an algorithmic solution or their solutions are too complex to be found. Currently, artificial neural network is used to solve problems that are unsolvable using logical systems. Our model has one input layer, two hidden layers, and one output layer as shown in Figure 10. The weights of the layers are optimized using an optimization technique during the training process of the weights.

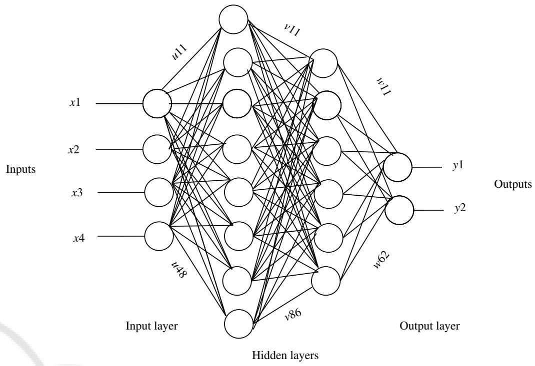


Figure 10: Deep forward neural network.

4.3 Auto Twitter Suicide Detector System (ATSDS)

The Limited-Memory BFGS (L-BFGS) is an optimization algorithm that approximates the Broyden–Fletcher–Goldfarb–Shanno (BFGS) algorithms using limited memory. The basic idea of L-BFGS is that it approximates a given objective function locally as a quadratic without calculating the second partial derivatives of the objective function. Thus, L-BFGS achieves faster convergence compared to the first-order optimization. It is a built in optimization algorithm in MLlib, and it has several parameters, such as Gradient, updater, numCorrections, maxNumIterations, regParam, and convergence tolerance .

ATSDS is built using L-BFGS and DFNN as shown in Figure 11. We study the effect of convergence tolerance on the performance in terms of accuracy of our proposed system. The convergence tolerance controls how much change is allowed when L-BFGS considered to converge. In our experiments, convergence tolerance is tuned to achieve the best accuracy of our proposed model.

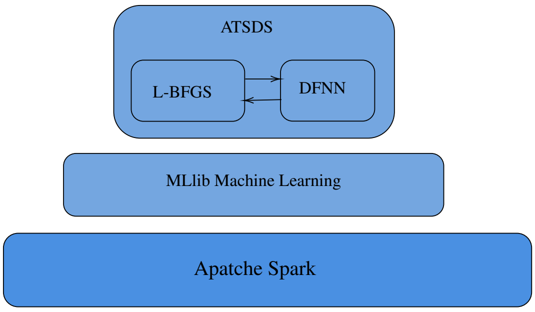


Figure 11: Auto Twitter Suicide Detector System (ATSDS).

5 EXPERIMENTATION AND RESULTS

In our model, the DFNN is set to have four layers: input layer, two hidden layer, and output layer. It is represented as DFNN (input layer, first hidden layer, second hidden layer, output layer). The input layer of the DFNN represents the number of features of the model, and it is set to 100 features. Each of the two hidden layers has 15 neurons. The output layer has two neurons, because the model has two output classes: suicide and non-suicide. The number of neuron in each layer is chosen after conducting preliminary experiments. In our experiments, we consider tuning the convergence tolerance as an important parameter in our optimization algorithm (L-BFGS).

The convergence tolerance is changed in the recommended range [0.001 : 0.15] when the number of neurons is fixed to 15 in each of the hidden layers. Then, we choose the best three convergence tolerance values (0.001, 0.01, 0.015) to study along with different number of neurons in hidden layers. The number of neurons in each hidden layer is changed in range [1 : 20] to achieve the best performance of our model. The maximum number of iterations in each case is set to 100,000 to create our ATSDS. The seed generator is set to 1234 in all experiments so that the results will be reproducible.

In order to evaluate the performance of our proposed system, we conducted several experiments that we report in this section. The accuracy is used to evaluate our system performance. The accuracy is the ratio of the number of correctly predicted instances to the total number of instances in the dataset. In our experiments, we study the effect of the convergence tolerance for L-BFGS and the number of neurons in the hidden layers over the accuracy of our proposed model.

5.1 Results and Discussion

In our experiments, the class label is whether it is suicide related tweet or non-related suicide tweet. Firstly, the model is trained on the entire dataset using two hidden layers of 15 neurons each, and then the model is tested on the same entire dataset. The result shows that the proposed model has high accuracy (96.7%). However, this system has bias because the dataset for training and testing are the same. Secondly, the dataset is then divided into two parts 75% training dataset and 25% testing dataset. The results show that the proposed model has relatively high accuracy 93.25%, but it is lower than the first case. In the second case, the model does not know any information about the testing dataset. However, when the train dataset is decreased to 25% and testing dataset is increased to 75%, the accuracy of the proposed model is decreased. In the third case, the cross validation technique is implemented for building our ATSDS, which is the most recommended technique to avoid bias in the model. The number of neurons in each hidden layer is tuned to achieve higher accuracy. The results show that the convergence tolerance of our optimization algorithm (L-BFGS) and the number of neurons in the hidden layers have significant effect on the performance of our proposed model.

The effect of convergence tolerance and the number of neurons in each hidden layer are tested to produce a high accuracy model. The result shows that number of hidden layers do not have that much effect on the accuracy of the model. The best number of neuron in the hidden layers is 15 as shown in Figure 12 and Table 1. Figure 12 shows the DFNN (100, 15, 15, 2), which has four layers. The input layer has 100 features; output layer has two classes suicide and non suicide; each of the two hidden layers has 15 neurons. The model has been tested on different convergence tolerance values. The results show that the best value of the convergence tolerance is 0.015.

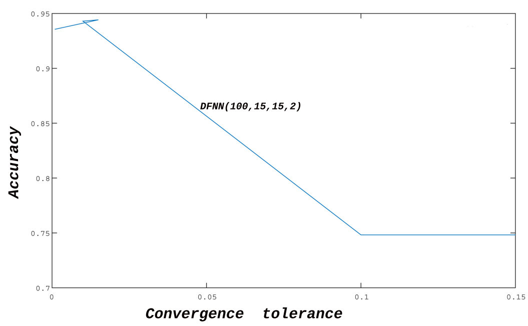


Figure 12: Accuracy vs convergence tolerance.

Figure 13 shows three different convergence values along with different number of neurons in the hid-

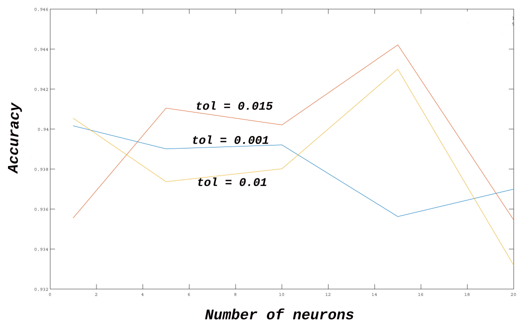


Figure 13: Accuracy vs number of neurons.

den layers in range [1 : 20]. The results show that the best convergence value is 0.015 along with different number of neurons. However, decreasing the number of neurons in both hidden layers below four neurons or increasing the number of neurons above 18 neurons deteriorates the accuracy of the model as shown the red line in Figure 13. Thus, when convergence tolerance is set to 0.015 and the number of neurons of both hidden layers are selected (5, 10, or 15) neurons, the ATSDS achieves better accuracy. The results show that when the convergence tolerance is increased to (0.1 or 0.15), the accuracy of the system is deteriorated as shown in Table 1.

Table 1: The accuracy and Convergence tolerance.

DFNN	Conv. tolerance	Accuracy
(100,15,15,2)	0.001	0.9356235
(100,15,15,2)	0.015	0.9442099
(100,15,15,2)	0.01	0.9429968
(100,15,15,2)	0.1	0.7482279
(100,15,15,2)	0.15	0.7482279

Table 1 shows that the best convergence tolerance values are (0.001, 0.015, and 0.01). Table 2 shows the accuracy of our proposed model along with changing those best convergence tolerance values and the number of neurons in both hidden layers. The output systems have high accuracy between 93.31% and 94.42% regardless of the number of neurons in the hidden layers.

6 CONCLUSION

The paper introduces a high accuracy Auto Twitter Suicide Detector (ATSD) system to auto detect users who are in danger with such kind of destructive suicide thoughts. ATSD is built on multi-purpose cluster computing system (Spark) using deep feed forward neural network and L-BFGS. The ATSD system is analyzed to choose the optimal parameters in terms of

Table 2: The accuracy and Convergence tolerance.

DFNN	Conv. tolerance	Accuracy
(100,1,1,2)	0.001	0.940160
(100,1,1,2)	0.015	0.935562
(100,1,1,2)	0.01	0.940536
(100,5,5,2)	0.001	0.939008
(100,5,5,2)	0.015	0.941045
(100,5,5,2)	0.01	0.937366
(100,10,10,2)	0.001	0.939207
(100,10,10,2)	0.015	0.940205
(100,10,10,2)	0.01	0.938012
(100,15,15,2)	0.001	0.935623
(100,15,15,2)	0.015	**0.944209**
(100,15,15,2)	0.01	0.942996
(100,20,20,2)	0.001	0.936995
(100,20,20,2)	0.015	0.935451
(100,20,20,2)	0.01	0.933187

the number of neurons and convergence tolerance values that increase the accuracy of the system. The proposed system is evaluated using Twitter dataset and achieved high accuracy (94.42%). We anticipate that our auto detector system can produce reliable results for Twitter posts, allowing authorities to mitigate the risk of suicide thoughts the threat our society. As future work, we intend to run more experiments on different datasets using different classifier to compare their accuracy.

ACKNOWLEDGEMENT

We thank Dr Nathaniel Osgood, University of Saskatchewan, Canada for his support with the Twitter dataset.

REFERENCES

Abboute, A., Boudjeriou, Y., Entringer, G., Azé, J., Bringay, S., and Poncelet, P. (2014). Mining twitter for suicide prevention. In *International Conference on Applications of Natural Language to Data Bases/Information Systems*, pages 250–253. Springer.

Becker, K. and Schmidt, M. H. (2005). When kids seek help on-line: Internet chat rooms and suicide. *Reclaiming Children and Youth*, 13(4):229.

Brown, L. M., Bongar, B., and Cleary, K. M. (2004). A profile of psychologists' views of critical risk factors for completed suicide in older adults. *Professional Psychology: Research and Practice*, 35(1):90.

Buchanan, D., Tourigny-Rivard, M., Cappeliez, P., Frank, C., Janikowski, P., Spanjevic, L., Malach, F., Mokry, J., Flint, A., and Herrmann, N. (2006). National guidelines for seniors' mental health: the assessment and treatment of depression. *Canadian Journal of Geriatrics*, 9(supplement 2):S52–S58.

Carlson, N. (2011). The real history of twitter. business insider. *Featured Articles From The Business Insider*, 13.

Conwell, Y., Pearson, J., and DeRenzo, E. G. (1996). Indirect self-destructive behavior among elderly patients in nursing homes: a research agenda. *The American Journal of Geriatric Psychiatry*, 4(2):152–163.

Glass Jr, J. C. and Reed, S. E. (1993). To live or die: A look at elderly suicide. *Educational Gerontology: An International Quarterly*, 19(8):767–778.

Gunn, J. F. and Lester, D. (2015). Twitter postings and suicide: An analysis of the postings of a fatal suicide in the 24 hours prior to death. *Suicidologi*, 17(3).

Hwang, J.-N. and Hu, Y. H. (2001). *Handbook of neural network signal processing*. CRC press.

Ikunaga, A., Nath, S. R., and Skinner, K. A. (2013). Internet suicide in japan: A qualitative content analysis of a suicide bulletin board. *Transcultural psychiatry*, 50(2):280–302.

Jashinsky, J., Burton, S. H., Hanson, C. L., West, J., Giraud-Carrier, C., Barnes, M. D., and Argyle, T. (2014). Tracking suicide risk factors through twitter in the us. *Crisis: The Journal of Crisis Intervention and Suicide Prevention*, 35(1):51.

Kaplan, A. M. and Haenlein, M. (2010). Users of the world, unite! the challenges and opportunities of social media. *Business horizons*, 53(1):59–68.

Karau, H., Konwinski, A., Wendell, P., and Zaharia, M. (2015). *Learning spark: lightning-fast big data analysis*. " O'Reilly Media, Inc.".

Matykiewicz, P., Duch, W., and Pestian, J. (2009). Clustering semantic spaces of suicide notes and newsgroups articles. In *Proceedings of the Workshop on Current Trends in Biomedical Natural Language Processing*, pages 179–184. Association for Computational Linguistics.

McCulloch, W. S. and Pitts, W. (1943). A logical calculus of the ideas immanent in nervous activity. *The bulletin of mathematical biophysics*, 5(4):115–133.

Navaneelan, T. (2012). *Suicide rates: An overview*. Statistics Canada Ottawa, Canada.

Pang, B., Lee, L., et al. (2008). Opinion mining and sentiment analysis. *Foundations and Trends® in Information Retrieval*, 2(1–2):1–135.

Pennebaker, J. W., Francis, M. E., and Booth, R. J. (2001). Linguistic inquiry and word count: Liwc 2001. *Mahway: Lawrence Erlbaum Associates*, 71(2001):2001.

Poulin, C., Shiner, B., Thompson, P., Vepstas, L., Young-Xu, Y., Goertzel, B., Watts, B., Flashman, L., and McAllister, T. (2014). Predicting the risk of suicide by analyzing the text of clinical notes. *PloS one*, 9(1):e85733.

Rosenblatt, F. (1958). The perceptron: a probabilistic model for information storage and organization in the brain. *Psychological review*, 65(6):386.

Sueki, H. (2015). The association of suicide-related twitter use with suicidal behaviour: a cross-sectional study of young internet users in japan. *Journal of affective disorders*, 170:155–160.

Won, H.-H., Myung, W., Song, G.-Y., Lee, W.-H., Kim, J.-W., Carroll, B. J., and Kim, D. K. (2013). Predicting national suicide numbers with social media data. *PloS one*, 8(4):e61809.

Yip, P. S., Chi, I., Chiu, H., Chi Wai, K., Conwell, Y., and Caine, E. (2003). A prevalence study of suicide ideation among older adults in hong kong sar. *International journal of geriatric psychiatry*, 18(11):1056–1062.