A Lexical Network Approach for Identifying Suicidal Ideation in Clinical Interview Transcripts

Ulya Bayram[1], Ali A. Minai[1], and John Pestian[2]

[1] Department of Electrical Engineering and Computer Science, University of Cincinnati, Cincinnati OH 45221, USA,
bayramua@mail.uc.edu, Ali.Minai@uc.edu
[2] University of Cincinnati Department of Pediatrics and Cincinnati Children's Hospital Medical Center, 3333 Burnet Avenue, Cincinnati, Ohio 45229, USA,
john.pestian@cchmc.org

Abstract. Preventing suicide requires early identification of suicidal ideation. In this research, we propose an approach to evaluate whether an individual's statements during a clinical interview can be classified as coming from a suicidal or non-suicidal mindset. To do so, we compare the statements with distinct lexical associative networks constructed from corpora of suicidal and control texts. Each node in these networks is a word, and the weight of the edge between every word pair indicates how strongly the words are associated in that corpus. Several metrics of association are evaluated in this work. Preliminary results show good classification performance with above 75% accuracy on novel test data.

Keywords: lexical networks, suicidal ideation, suicide, clinical interviews, word association, spreading activation, classification

1 Introduction

Worldwide, there are approximately 800,000 suicides a year, and at least six people are devastated by each of these deaths [15]. In the United States, suicide is among the top ten leading causes of death [10]. Most methods for detecting suicidal risk rely on the identification of suicidal ideation via self-reported answers to direct questions [5, 23], which can be inaccurate [26] because individuals often try to hide suicidal thoughts [29].

This has motivated the search for other clues, such as objective suicidal thought, or behavioral markers [29]. Various data collections obtained from individuals with suicidal ideation have allowed researchers to use common supervised machine learning methods and experiment with some standard natural language processing (NLP) techniques. Some of these studies have shown state-of-the-art classification rates within their datasets, but their real-world application performance is yet to be determined. Many of these approaches have a common drawback: they do not account for the cognitive aspect of the problem which is

a key element in predicting suicide [28]. In this study, we describe an approach that accounts for this cognitive element.

During self-expression, thoughts and ideas are organized into sentences [14], and the use of specific words in the same sentence is a strong indicator that the corresponding concepts are associated in the individual's mind. We use this to build *lexical networks* from interview transcripts to capture the pattern of association between concepts in the suicidal and non-suicidal mindsets and determine which mindset a given expression is more consistent with. Experimental results show that the method detects suicidal ideation with greater than 75% accuracy for novel expressions, i.e. those not used to build the reference lexical networks. As a followup, we also experiment with spreading activation [1, 33] in the lexical network to further enrich the discriminative process. This allows expressions to be evaluated not only based on their content but also on related thoughts. Results show some initial promise for improving the potential for identifying suicidal ideation, though this remains to be validated. The approach we describe is not restricted by the suicidal ideation detection problem, and can be applied to any other binary text-classification problem where reference corpora are available.

2 Method

2.1 Data

This study uses two prospectively assembled data collections containing interviews conducted with volunteering subjects, all native English speakers, by a mental health professional asking five simple questions regarding the subject having any fears, feeling hurt, having hopes, feeling angry, and having secrets. The first collection includes 60 transcripts of interviews conducted with adolescents [30], where half are patients admitted to the Cincinnati Children's Hospital Medical Center (CCHMC) for suicidal ideation or attempts and the rest are the patients admitted for orthopedic injuries with no recent or past suicidal ideation. The second data collection extends the first. It includes interviews collected from patients admitted to one of three hospitals: CCHMC, University of Cincinnati Medical Center (UC), and Princeton Community Hospital (PCH) who were either suicidal, mentally ill, or a member of the control group [32]. In contrast with the first collection, this study includes follow-up interviews conducted a month later, mentally ill patients, and adults. Since we are interested in detecting suicidal ideation, we exclude the mentally ill patient interviews; we merge the two collections and get a final dataset with 470 transcribed and annotated interviews. To standardize data for analysis, we use standard NLP techniques: converting each character to lowercase, removing punctuation while keeping only the end-of-sentence information, removing the English language stop words available in the Natural Language Toolkit (NLTK) [7], and also removing infrequent words with fewer than 5 occurrences.

2.2 Lexical Networks

Building networks from words and word associations is one approach for semantic memory research [1, 9], and in research problems aiming to discover the semantic relatedness between words in the lexical domain [27, 34, 8, 11, 16, 19, 12, 17]. In this study, we use this approach to identify the thought-patterns models of suicidal and non-suicidal individuals. This method is based on the idea that thoughts emerge from a substrate of associations between concepts [21, 18, 20, 6], and that the analysis of associations between words in the statements of an individual can provide access to the semantics of their thinking [25, 4, 12, 16, 17].

We initialize both lexical networks (suicidal and control) using the respective training corpora by adding the pre-processed words spoken by suicidal ($V_{suicidal}$) and non-suicidal ($V_{control}$) interviewees as nodes. Since thoughts are organized and expressed through sentences [14], we use sentence-level word associations to establish connections between the words in the networks. However, NLP literature contains various word association metrics [11, 12, 22]. In our approach, we experiment with three of such metrics. The joint probability metric $p(i, j)$ represents the probability of observing words i and j together in the same sentence in the training corpus. The correlation coefficient metric, $cc(i, j)$, is a normalized covariance measure in the occurrence of i and j, and uses the in-sentence occurrence probabilities (p_i, p_j) of the words (see Eq. 1). Finally, the point-wise mutual information metric (PMI), $pmi(i, j)$, measures the statistical dependence in the occurrence of i and j (see Eq. 2). We use a normalized version of the PMI metric.

$$cc(i,j) = \frac{p(i,j) - p_i p_j}{\sqrt{p_i(1-p_i)p_j(1-p_j)}} \tag{1}$$

$$pmi(i,j) = log\big(\frac{p_i p_j}{p(i,j)}\big)/log(p(i,j)) \tag{2}$$

Following the computation of word associations from training corpora, we set the negative values of the correlation coefficient and PMI meaning non-association to zero weights. Thus, non-associated word pairs remain unconnected in the networks and the positive edge weights indicate the connection strengths between words. Samples from the two networks are shown in Fig. 1, where the differences between the suicidal and control networks are visually evident.

2.3 Evaluating The Fitness of Thoughts

To determine whether a given novel statement is more consistent with a suicidal or non-suicidal mindset, we take the words from the statement and extract the sub-networks of each reference lexical network with only these words. An example of the sub-network obtained by matching the words in a short test statement with those in the suicidal lexical network with joint probabilities as edge weights is shown in Fig. 2a. We hypothesize that the strength of connectivity within the two sub-networks obtained in this way can identify which lexical network the test statement fits better in, thus determining whether the underlying statement is more consistent with a suicidal mindset or a non-suicidal one. To do this, we

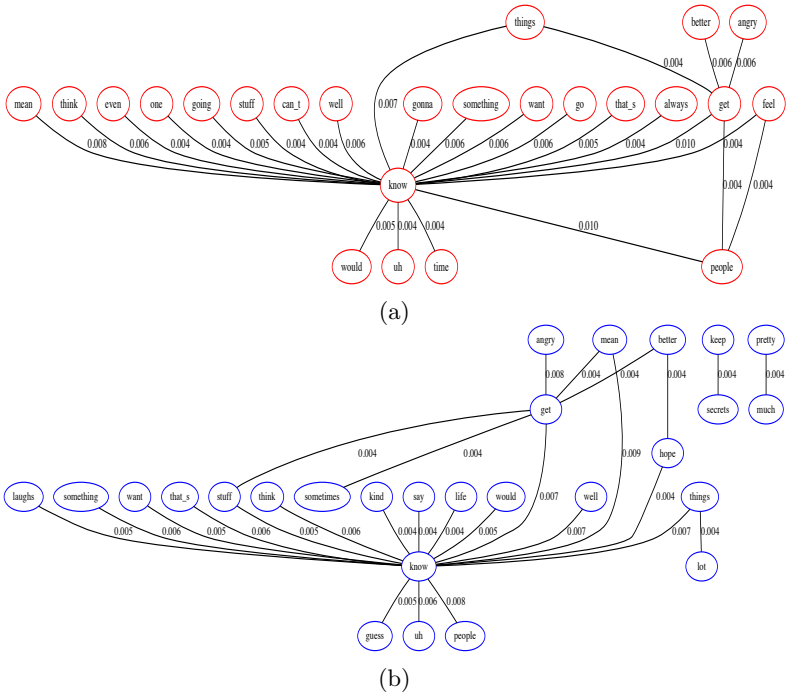


Fig. 1. The strongest connections in joint probability networks of suicidal and control training corpora obtained by thresholding the edge weights, returned (a) out of 48, 169 edges in the suicidal network, (b) out of 38, 132 edges in the control network.

define a *fitness function* that is applied to the extracted sub-network, E:

$$\Phi(E) = d_e^S(E) - d_e^C(E) \tag{3}$$

$$d_e^G(E) = d^E - d^G = \frac{\sum_{i,j \in E} w_{i,j}}{n_E(n_E-1)/2} - \frac{\sum_{i,j \in G} w_{i,j}}{n(n-1)/2} \tag{4}$$

where S and C superscripts are the suicidal and control networks, G is the reference lexical network (suicidal or control), n is the number of nodes in G, n_E is the number of nodes in sub-network E, and $d_e^G(E)$ is the *excess weight density* of E with respect to G, which quantifies the degree to which the strength of connectivity within E exceeds the value that would be expected for a random sub-network of the same size extracted from G. Such measures are commonly used in network studies [13, 24].

2.4 Spreading Activation

Spreading activation (SA) theory was developed by Quillan [33] as an attempt to model the associative nature of human semantic memory. Although there are many spreading activation methods [1, 2, 3, 31], the main idea is based on assigning initial activation values to nodes, and by some rules, letting the activation spread to the rest of the network via associative edges. To limit spreading, all

spreading activation approaches include a loss factor that can be a decay function or a simple threshold parameter. Spreading activation is a directed process whereby active nodes activate those to which they are connected with sufficiently strong edges.

When a test expression is presented to the system, nodes corresponding to its words are activated, and this activation is allowed to spread until the limiting process stops the activation. All nodes activated in this episode then become part of the sub-network that is tested for fitness.

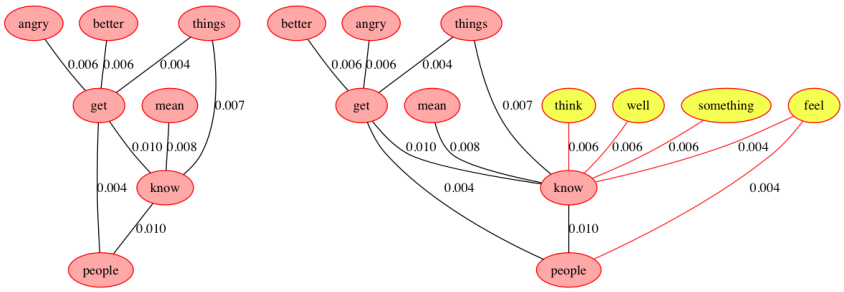


Fig. 2. Sample sub-networks obtained (a) without using spreading activation, (b) using spreading activation, where the nodes and connections added by spreading are highlighted.

As a result of the SA process, we obtain two sub-networks comprising the initial sub-networks plus words and connections activated by the spreading process. Thus, these sub-networks can be considered enriched or augmented versions of the original ones. Sample sub-networks obtained without SA and with it are shown in Fig. 2a and Fig. 2b respectively, with nodes added by the spreading activation results highlighted. The fitness of the augmented sub-networks is evaluated as with the original ones.

3 Results

For experiments, we increased the effective number of training and test cases by randomly and independently choosing ten test-training sets in $1/3 - 2/3$ ratio from both the suicidal and control corpora, training and testing separately on each set, and aggregating the results, giving a total of 750 test data points for each corpus.

The first experiment evaluated the performance of the method without spreading activation. A test data item was labeled as suicidal or control based on which lexical network gave it a higher fitness. The results obtained from $1,500$ test cases are summarized in Table 1, where the best association metric is the correlation coefficient, with above 75% accuracy. Having such a high accuracy in this difficult task indicates that lexical networks are a promising method for doing such evaluations.

Association metric	Accuracy	AUC score ± stdev
Joint probability	64.3%	74.8 % ± 2.5%
Point-wise Mutual Information (PMI)	73.6%	79.3% ± 2.3%
Correlation coefficient	**75.6%**	**81.6%** ± 2.1%

Table 1. Suicidal ideation detection results measured by classification accuracy and area under the ROC curve (AUC) obtained using three association metrics on 1,500 test cases.

The second experiment uses spreading activation (SA) to evaluate the effect of the augmented sub-networks on detecting suicidal ideation. The addition of more words and connections to the sub-networks by spreading activation results an accuracy of 74.5% using the lexical networks built by the correlation coefficient metric, which is actually a very small *decrease* in accuracy over the original method. However, the potential benefits of the approach become clearer by looking at the number of true positives (TP) and false positives (FP), where positive indicates the suicidal case. Comparing TP and FP values between results obtained with (TP=555, FP=187) and without spreading activation (TP=550, FP=166) shows that having spreading activation allows an increase in TP rates, but at the cost of an increase in control cases being identified as suicidal. In this life-critical application, however, this may well be an acceptable trade-off.

4 Conclusion

In this study, we have proposed a lexical network approach to detecting suicidal ideation in self-expression data. Experiments on a novel dataset with actual interviews have shown that the approach is able to classify suicidal and non-suicidal ideation with high accuracy, which demonstrates the success of the approach not only for this application, but also for other binary classification problems involving textual data. The use of spreading activation as a part of the approach is potentially useful because associations play a key role in the organization of thought. However, the benefit remains to be demonstrated in the specific case of suicidal ideation analysis. Experiments have provided evidence that using spreading activation can improve suicidal detection rates at the cost of including more false positives. Since identifying suicidal individuals is far more important than potentially misidentifying some non-suicidal individuals as suicidal, the benefits of spreading activation may outweigh its cost. Also, while the data used was based on a careful identification of suicidal and non-suicidal individuals, it could not possibly control for every expression. Even non-suicidal individuals can occasionally express ideas consistent with a suicidal mindset, and vice versa. Thus, classifying expressions as a way of identifying an individual's entire mindset is inherently subject to some degree of uncertainty, which can only be alleviated by using more extensive data from each individual.

The methods described in this study represent only an initial phase of our research. The fact that even the simple methods described above give a high degree of discrimination indicates that they can serve as the basis for more sophisticated methods, which we will report on in the future.

Acknowledgment

We are grateful for the support of the University of Cincinnati College of Medicine, and the Department of Electrical Engineering and Computer Science. We are also grateful for the support of the Cincinnati Children's Hospital Medical Center, and its division of Biomedical Informatics.

References

[1] Anderson JR (1983) A spreading activation theory of memory. Journal of Verbal Learning and Verbal Behavior 22(3):261–295
[2] Anderson JR (1996) Act: A simple theory of complex cognition. American Psychologist 51(4):355
[3] Anderson JR, Bothell D, Byrne MD, Douglass S, Lebiere C, Qin Y (2004) An integrated theory of the mind. Psychological Review 111(4):1036
[4] Bales ME, Johnson SB (2006) Graph theoretic modeling of large-scale semantic networks. Journal of Biomedical Informatics 39(4):451–464
[5] Beck AT, Kovacs M, Weissman A (1979) Assessment of suicidal intention: The scale for suicide ideation. Journal of Consulting and Clinical Psychology 47(2):343
[6] Benedek M, Jauk E, Fink A, Koschutnig K, Reishofer G, Ebner F, Neubauer AC (2014) To create or to recall? Neural mechanisms underlying the generation of creative new ideas. NeuroImage 88:125–133
[7] Bird S, Klein E, Loper E (2009) Natural Language Processing with Python. O'Reilly Media
[8] Budanitsky A, Hirst G (2006) Evaluating Wordnet-based measures of lexical semantic relatedness. Computational Linguistics 32(1):13–47
[9] Collins AM, Loftus EF (1975) A spreading-activation theory of semantic processing. Psychological Review 82(6):407
[10] Coppersmith G, Leary R, Whyne E, Wood T (2015) Quantifying suicidal ideation via language usage on social media. In: Joint Statistics Meetings Proceedings, Statistical Computing Section. JSM
[11] De Deyne S, Storms G (2008) Word associations: Network and semantic properties. Behavior Research Methods 40(1):213–231
[12] Doumit S, Marupaka N, Minai AA (2013) Thinking in prose and poetry: A semantic neural model. In: Proceedings of IJCNN 2013
[13] Frank KA (1995) Identifying cohesive subgroups. Social Networks 17(1):27–56
[14] Frege G (1956) The thought: A logical inquiry. Mind 65(259):289–311
[15] Geneva: World Health Organization (2017) Preventing suicide: A resource for media professionals. Tech. rep., World Health Organization
[16] Kenett Y, Kenett D, Ben-Jacob E, Faust M (2011) Global and local features of semantic networks: Evidence from the Hebrew mental lexicon. PLoS One 6:e23,912

[17] Kenett Y, Anaki D, Faust M (2014) Investigating the structure of semantic networks in low and high creative persons. Frontiers in Human Intelligence 8:Article 407
[18] Landauer T, Dumais S (1997) A solution to plato's problems: The latent semantic analysis theory of acquisition, induction and representation of knowledge. Psychological Review 104:211–240
[19] Marupaka N, Iyer LR, Minai AA (2012) Connectivity and thought: The influence of semantic network structure in a neurodynamical model of thinking. Neural Networks 32:147–158
[20] McRae K, de Sa V, Seidenberg M (1997) On the nature and scope of featural representations of word meaning. Journal of Experimental Psychology: General 126:99–130
[21] Mednick S (1962) The associative basis of the creative process. Psychological Review 69(3):220–232
[22] Mei M, Vanarase A, Minai AA (2014) Chunks of thought: Finding salient semantic structures in texts. In: Proceedings of IJCNN 2014
[23] Miller IW, Norman WH, Bishop SB, Dow MG (1986) The modified scale for suicidal ideation: Reliability and validity. Journal of Consulting and Clinical Psychology 54(5):724
[24] Moody J, White DR (2003) Structural cohesion and embeddedness: A hierarchical concept of social groups. American Sociological Review 68:103–127
[25] Motter AE, de Moura APS, Lai YC, Dasgupta P (2002) Topology of the conceptual network of language. Physical Review E 65:065,102(R)
[26] Murray D (2016) Is it time to abandon suicide risk assessment? British Journal of Psychiatry Open 2(1):e1–e2, DOI 10.1192/bjpo.bp.115.002071
[27] Nelson DL, Xu J (1995) Effects of implicit memory on explicit recall: Set size and word frequency effects. Psychological Research 57:203–214
[28] Neuringer C (1988) The thinking processes in suicidal women. Why Women Kill Themselves pp 43–52
[29] Nock MK, Park JM, Finn CT, Deliberto TL, Dour HJ, Banaji MR (2010) Measuring the suicidal mind: Implicit cognition predicts suicidal behavior. Psychological Science 21(4):511–517
[30] Pestian JP, Grupp-Phelan J, Bretonnel Cohen K, Meyers G, Richey LA, Matykiewicz P, Sorter MT (2015) A controlled trial using natural language processing to examine the language of suicidal adolescents in the emergency department. Suicide and Life-Threatening Behavior
[31] Pestian JP, Matykiewicz P, Duch W, Glauser TA, Kowatch RA, Grupp-Phelan JM, Sorter M (2016) Processing text with domain-specific spreading activation methods. US Patent 9,477,655
[32] Pestian JP, Sorter M, Connolly B, Bretonnel Cohen K, McCullumsmith C, Gee JT, Morency LP, Scherer S, Rohlfs L (2017) A machine learning approach to identifying the thought markers of suicidal subjects: A prospective multicenter trial. Suicide and Life-Threatening Behavior 47(1):112–121
[33] Quillan MR (1966) Semantic memory. Tech. rep., DTIC Document
[34] Steyvers M, Tenenbaum JB (2005) The large-scale structure of semantic networks: Statistical analyses and a model of semantic growth. Cognitive Science 29(1):41–78