ISSN 2319-2720
Volume 7, No.2, April – June 2018
International Journal of Computing, Communications and Networking
Available Online at http://www.warse.org/ijccn/static/pdf/file/ijccn27722018.pdf
https://doi.org/10.30534/ijccn/2018/27722018

# Assessing Antidepressants Using Intelligent Data Checking and Mining of Online Fora

**Prof. Jagadeesh BN[1], Manoj S[2], MD Nadeem [3], Allahrakha Shaikh[4], Jamir Sk[5]**
[1]Assistant Professor EWIT, India,jagadeeshbn.cse@ewit.edu
[2]Student, EWIT, India, manurokz006@gmail.com
[3]Student, EWIT, India, mohammednadeemhusain79@gmail.com
[4]Student, EWIT, India,allahrakha10205@gmail.com
[5]Student, EWIT, India,jamir13073@gmail.com

## ABSTRACT

Depression is a worldwide wellbeing concern. Informal organizations enable the influenced populace to share their encounters. These encounters, whenever mined, extricated, and broke down, can be changed over into either notices to review drugs (perilous reactions), or benefit change (mediations, treatment choices) in view of perceptions got from client conduct in sadness related social networks. Our aim was to build up a weighted system model to speak to client action on social wellbeing systems. This empowered us to precisely speak to client collaborations by depending on the information's semantic substance. Our three-advance strategy utilizes the weighted system model to speak to client's movement, and system bunching furthermore, module investigation to describe client communications and concentrate facilitate learning from client's posts. The system's topological properties reflect client movement, for example, posts' general point also as timing, while weighted edges mirror the posts semantic substance what's more, similitudes among posts. The outcome, an amalgamation from word information recurrence, measurable investigation of module content, and the demonstrated wellbeing system's properties, has enabled us to pick up knowledge into customer estimation of antidepressants. This approach will permit all gatherings to take an interest in enhancing future wellbeing arrangements of patients experiencing sadness.

**Keywords:Datamining, depression, network analysis, online fora, semantic analysis, social media, user sentiment.**

## I. INTRODUCTION

Melancholy is the main source of incapacity around the world,a noteworthy supporter of the worldwide weight of ailment, and is influencing in excess of 350 million individuals overall [1], [2]. Untreated sadness has been connected to issues extending from stroke to coronary conduit malady [3], two of the main ten driving reasons for death on the planet in 2012 [4]. Not as much as half of the worldwide populace gets the best possible consideration and treatment in both the immature nations and industrialized nations [5]. Gatherings and online networking sites devoted to dejection have as of late jumped up for patients and human services specialists to share their encounters from overseeing gloom in their every day lives to their responses to antidepressants. Such voluminous data can give boundless chances to patients, medicinal services associations, and industry to enhance arrangements through canny information mining, extraction, and examination.

A web-based social networking system is a virtual systems administration condition that is made out of hubs and edges. Its substance can be demonstrated Furthermore, separated utilizing computational apparatuses that can outline, detail expectations, and evaluate client connections. Graphical portrayal can outwardly speak to the data. A sociomatrix can speak to an informal community's structure. Topological parameters, for example, hub degrees and system densities can clarify particular elements inside a system and particular calculations can outline data rich structures (for example, bunches). Distinguishing these groups empowers hub (or on the other hand bunch) focused data mining. Such urgent information can help social insurance associations, doctors, staff, and patients to enhance administrations in light of input from "savvy" information mining of wellbeing particular online networking destinations.

Our present examination includes more semantic setting to our system display by considering, in the demonstrate building step, data contained inside the posts. We added weights to the system edges that mirror the semantic similitude between the edge-associated hubs. The edge weights themselves result from an underlying preprocessing step utilizing the k-implies grouping. We enhanced the examination of the recovered organize modules by utilizing measurable testing, construct basically with respect to the hypergeometric test, which empowered us to discover essentially overrepresented terms in specific modules. We additionally accounted for how the client positions and their reaction time to any given post on particular data influenced the online networking system's inner dynamics.We at that point distinguished persuasive clients utilizing the hyperlink-actuated theme look (HITS) calculation.

## II. METHODS

### A. Initial Data Search and Collection

The initial step was to look for the most looked for after gatherings devoted to melancholy. Our last rundown, which yielded the accompanying .Thus, depressionforums.org was our choice.

We chose the forumDepressionforums.org as our main source of information based on our findings on the first list (the numberof users, posts, and threads).

**B. Text Mining, Preprocessing, and Tagging**

Adata gathering, investigation, and preparing treewas created in Rapidminer (www.rapidminer.com) to find the most incessant words (positive, negative, and symptoms) to discover their term-recurrence backwards record recurrence (TF-IDF) scores inside each post. Fig. 2.1 demonstrates the information gathering and handling tree. The dataset was transferred ("Read Excel"), handled ("Process Documents to Data") utilizing subcomponents ("Extract Content," "Tokenize," "Change Cases," "Channel Stopwords," "Channel Tokens," individually) that separated overabundance clamor (incorrectly spelled words, normal stop words, and so on.) to guarantee quantifiable variable consistency. The outcome ("Processed Data") contained the last word list, with each word containing a particular TF-IDF score.

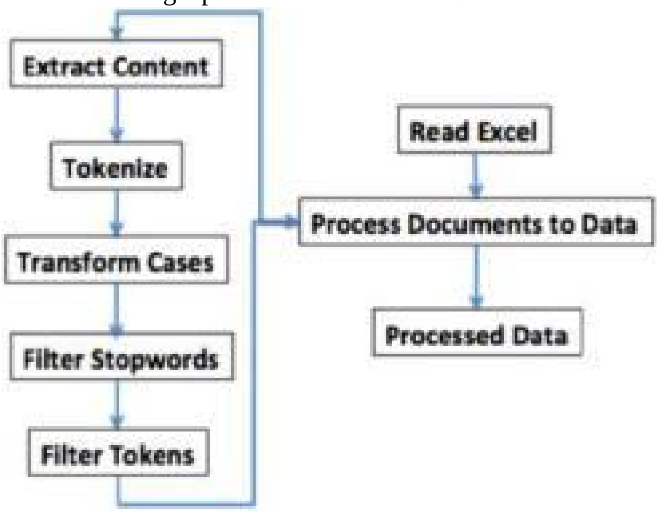


Fig. 2.1. Processing tree in Rapidminer to ascertain the TF-IDF scores of words in the data

We assigned every word a specific score using the following formula:

$$\text{score}_{t,d} = \begin{cases} (\log tf_{t,d} + 1)\log\dfrac{n}{x_i} & \text{if } tf_{t,d} \geq 1 \\ 0 & \end{cases} \quad . \qquad (1)$$

.

where tfi,d is word frequency (t) in the document (d), n is the sum of documents in the entire collection, and xt is the number of documents where t is present. TF-IDF is a widely used standard frequency measure [42].

**C. Semantic Similarity of User Postings**

The TF-IDF scores in each post were assembled in view of a delegate word set present all through the discussion and reflects the posts' semantic substance. In this way, we saw a TF-IDF vector as the semantic profile of each post. Subsequently, different measures of likenesses can be inferred to reflect how close the semantic profiles of two posts are, e.g., Euclidian separation or relationship. Moreover, bunching examination can be performed to recognize gatherings of comparative semantic profiles.We utilized k-implies bunching [48] to generally amass the semantic profiles of all posts from our gathering, as a preprocessing step fundamental for the organize based displaying.

**D. Network-Based Modeling of Forum Postings**

Forum postingmovement comprising of strings containing thousands of postings and answers were demonstrated into a substantial user centric organize. The displaying approach went for reflecting client connections while at the same time thinking about the posts' semantic substance. The hubs in our system compare to gathering clients and interfacing coordinated edges compare to two distinctive sorts of collaborations: direct and setting associations. Coordinate communications compare to guide client to-client answers utilizing the gathering's "Answer" choice. These connections were displayed with bidirectional edges associating the two relating hubs. This enabled us to show the common trade of data between a publication and a direct replier.Fig. 1 shows an example of our thread-based network modelling approach.

In Fig. 2.2, Node #1 is the string initiator and all things considered there are directional edges connecting this hub to every single other hub inside this particular string. Hub #5 is an immediate answer to the string initiator and in that capacity it is linkedwith a bidirectional association with the string initiator. So also, Node #3 is specifically answering to Hub #2, reflected by the bidirectional edge connecting the particular hubs.

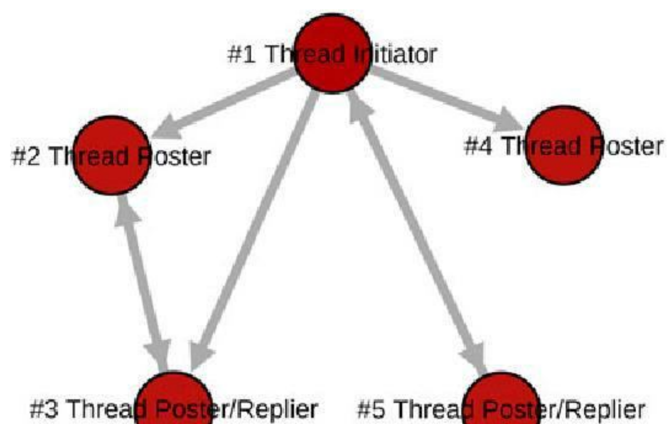


Fig. 2.2, Network model we considered: nodes represent users/posts and the edges represent information transferred among users.

**E. Identifying Subgraphs (Network Modules)**

Our modeling frameworkhas therefore changed over the gathering posts into a huge directional weighted system containing a number of thickly associated units (or modules) [see Fig. 2.3(b)]. The succession is as per the following. A. Dialect preparing piece. To begin with, the posts gathered from the discussion by means of Rapidminer are preprocessed utilizing the NTLK Toolbox and GATE (Step A1) and changed into a wordlist (Step A2). At this progression, coordinate mapping to the UMLS labeling and synonymous thesaurus is utilized to distinguish words speaking to medicinal terms and depressionrelated reactions and to Thesaurus Synonyms Database and Merriam-Webster for equivalent words coordinating. In view of the two wordlists, gathering posts are changed into numerical vectors containingword recurrence based TF-IDF scores, which are thusly bunched utilizing the k-implies technique. Also, a database comprising of all wordlist terms found in each post is made (advance A3).B. System preparing square. In parallel, gathering posts and answers are displayed as a weighted coordinated system (Step B1).

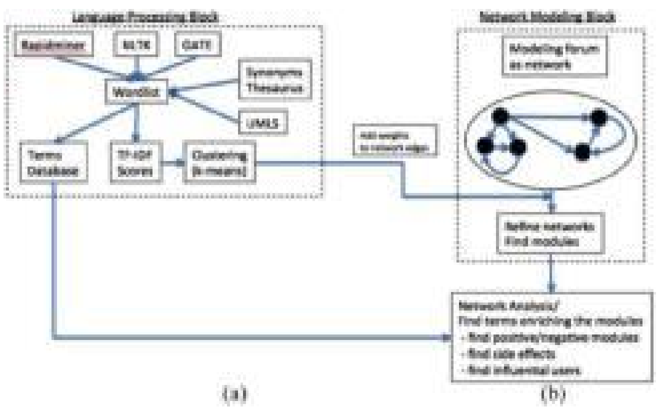


Fig. 2.3. Diagram describing the framework of our network-based analysis.

**F. Modules' Terms Enrichment Analysis**

We at that point continued to discover terms that are essentially overrepresented inside every one of the modules found by the networkpartitioning calculation (see Section IIE) utilizing a database containing the terms clarifying each post (see Section II-B) what's more, the hypergeometric improvement test that depends on the hypergeometric conveyance [54].

**G. Network-Based Identification of Influential Users**

We continued to distinguish compelling clients in the recovered modules. To this objective, we utilized the HITS, which is a technique at first produced for site pages interface examination [55], [56]. It recovers, through shared recursion, two numbers to a hub: an specialist and a center score. A hub has higher specialist when it gets approaching edges from hubs with higher center point scores. A higher center point score happens when a hub has active edges to numerous high specialist hubs.The calculation finds the expert what's more, center scores of every hub in a module by means of a progression of emphasess comprising of two stages: expert refresh (amid which each hub's power score is processed as the aggregate of all center scores of every hub to which it is associated through approaching edges) also, center point refresh (duringwhich every hub's center score is figured as the entirety of the specialist scores of every hub to which is sends

active edges). Thusly, the specialist and center point score for each hub is figured as takes after: 1) introduce the hubs specialist furthermore, center scores to 1; 2) execute the expert refresh step; 3) execute the center point refresh control; and 4) standardize the estimations of the pecialist (center point) scores by partitioning every expert (center) score by square base of the whole of squares of all specialist (center) scores [55]. Data scatters from definitive hubs. Center points interface to legitimate hubs, and along these lines, they dealer data streaminside the system. This approach for distinguishing powerful clients has the upside of considering both the system's basic properties, and the directionality of data stream.

## III. PROPOSED SYSTEM

Our aim is to develop a weighted network model to represent user activity on social health networks.This enabled us to accurately represent user interactions by relying on data's semantic content.Our three step method uses the weighted network model to represent users activity, and network clustering and module analysis to characterize user interactions and extract further knowledge from user's post.

The networks topological properties reflect user activity such as post's general topic as well as timing, while weighted edges reflect the posts semantic content and similarities among postsActivity.

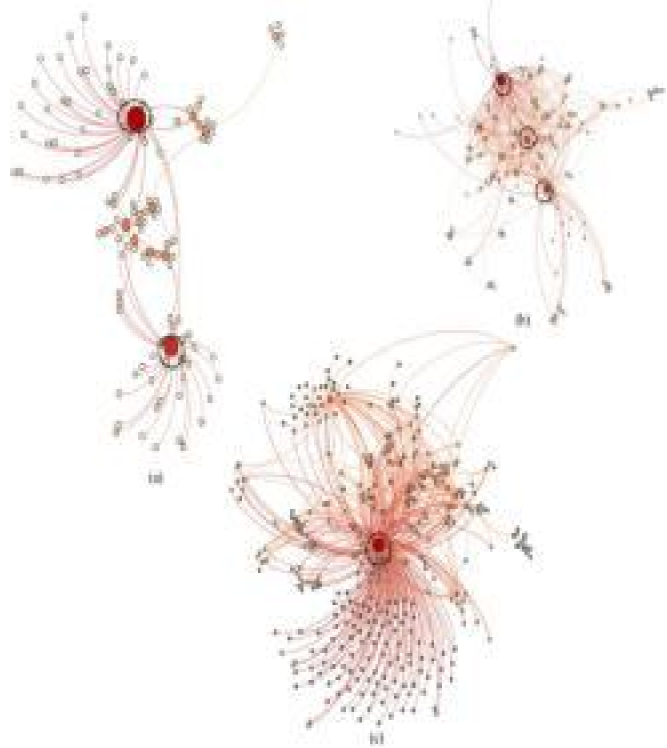

Fig. 3.1. Modules retrieved after the network's partitioning. Circled nodes represent influential users identified as described in Section IIG.

## IV. RESULTS

We gathered information from the most well known sadness related message board: Depressionforums.org. A sum of 7 726 posts (posted between July 2004 and October 2014, with a normal of ~2/day) were consequently downloaded and preprocessed utilizing RapidMiner, NLTK, and TAKE as

depicted in Sections II-An and II-B. Tokenized, sifted and labeled terms were at that point pruned utilizing the equivalent words databases, and after that, mapped to the UMLS. This brought about a rundown of 277 terms that passed a edge of n > 10 appearances in the posts and signified side impacts, drugs, and additionally positive and negative terms. Our information were then changed into a numerical framework (7726 × 277)
containing the TF-IDF scores for all gathering posts. For the kmeans investigation step, we utilized k = 20 groups for the underlying harsh bunching of the TF-IDF inferred semantic profiles. This esteem was dictated by finding the base estimation of theDavis– Bouldin file, which compares to an ideal grouping [58]. In (1), η1 and η2 were picked as 0.4 and 0.6, separately. Our system displaying approach yielded an underlying freely associated organize, connecting all clients inside the discussion. Ensuing module distinguishing proof utilizing the strategies portrayed in Area II-E yielded an ideal system dividing containing 14 thickly associated modules. We shifted our scale parameter inside the interim t [0, 2] of every 0.1 additions, as proposed in [58] and [59]. Fluctuating the scale parameter brought about a set of allotments running from modules containing single person clients (for scale parameter t = 0), to extensive modules (for estimations of t near the maximum furthest reaches of the interim). The ideal parcel (expanding the soundness based quality measure) was gotten for t = 1. Fig.3.1 indicates three of the 14 modules recognized. When modules were distinguished, we additionally described the modules content utilizing the term enhancement investigation portrayed in Section II-F. Supplementary Table I displays the terms distinguished to be altogether advancing the modules. We characterized a measure that enabled us to measure the proportion of positive and negative termsthat enrich the modules to provide estimate on the user's generalopinion within a module

ri=(pi/pt)/(ni/nt) (2)

where $p_i$ is the number of positive terms enriching module *i*, $p_t$ is the total number of positive terms in the wordlist, $n_i$ is thenumber of negative terms enriching module *i*, and $n_t$ is the totalnumber of negative terms in the wordlist.
In view of this measure, we portrayed modules to express predominately positive mind-set/assessments when their comparing ri measure was more noteworthy than 0.5 and dominatingly negative at the point when ri < 0.5. Out of the 14 modules recognized, 5 have a place with the positive class, while 6 modules have a place with the negative class. The staying three modules were not improved with not one or the other negative, nor positive terms. Strangely, we watched that positive class modules had essentially higher normal client positions *(3.62 ± 0.19*) than negative class modules (2.66 ± 0.24), p < 0.01 (Using Student's t-test—FDR amended). Client positions were numerically coded, going from 0—comparing to newusers to 13—comparing to premium platinum individuals (most elevated positioned clients). Discussion rankings depend on clients' action on the discussion and the quantity of aggregate postings.We additionally watched essentially higher normal hub degrees, and also weighted hub degrees (p < 0.01) . This specifically mirrors the actuality that clients from positive class modules are all things considered associated to a greater number of clients than those in negative class modules. In addition, the way that edges interfacing clients in positive class modules have higher weights all things considered, mirror the way that posts in these modules are semantically more homogeneous. Our investigation additionally uncovered essentially improved reaction terms (p < 0.01) in three out of fourteen recovered modules. Modules 1 and 5 were essentially advanced for symptom term anorgasmia. These two modules were additionally fundamentally improved for drugs terms citalopram (Celexa), cyclobenzaprine (Flexeril), duloxetine (Cymbalta), lorazepam, and mirtazapine (Module 1) and citalopram, chlorpromazine, and venlor (Module 5). Module 8 was advanced for symptom hypertension as well as the medication term promethazine. Our outcomes affirm clinical
inquire about discoveries, which already announced anorgasmia as a normal symptom of citalopram, cyclobenzaprine, duloxetine, also, chlorpromazine [59]–[61]. Furthermore, hypertension was detailed as a symptom of promethazine [62].

## V. CONCLUSION

Our aim is to develop a weighted network model to represent user activity on social health networks.

This approach will allow all parties to participate in improving future health of patients suffering from depression.

To conclude, we believe the use of intelligent data mining tools is an opportunity to greatly improve the quality of healthcare by consumers, healthcare workers, and the industry while reducing costs.

## REFERENCES

[1] World Health Organization. (2015, Jan. 5).Depression Fact Sheet No. 369.[Online].Available: http://www.who.int/mediacentre/factsheets/fs369/en/

[2] A. J. Ferrari *et al.*, "Burden of depressive disorders by country, sex, age,and year: Findings from the global burden of disease study 2010," *PLOSMed.*, Nov. 2013
.
[3] WebMED. (2015, Jan. 6). Depression Health Center: Untreated Depression. [Online].Available: http://www.webmd.com/depression/guide/untreated depression-effects

[4] World Health Organizations. (2015, Jan. 7). The Top 10 CausesDeath.[Online].Available:who.int/mediacentre/factshe ets/fs310/en/

[5] Ibid.

[6] L. Getoor and C. Diehl, "Link mining: A survey," *SIGKDD Explor. Newslett.*, vol. 7, pp. 3—12, Dec. 2005.

[7] Q. Lu and L. Getoor, "Link-based classification," in *Proc. 20th Int. Conf.Mach. Learn.*, Washington, DC, USA, 2003, pp. 496–503

[8] A. Ng, A. Zheng, and M. Jordan, "Stable algorithms for link analysis,"in *Proc. SIGIR Conf. Inf. Retrieval*, New Orleans, LA, USA, 2001,pp. 258–266

[9] B. Taskar, M. Wong, P. Abbeel, and D. Koller, "Link prediction in relationaldata," presented at the *Advances Neural Information ProcessingSystems.*, Vancouver, BC, USA, 2003

[10] D. Liben-Nowell and J. M. Kleinberg, "The link prediction problem for social networks," *J. Amer. Soc. Inf. Sci. Technol.*, vol. 57, pp. 556–559,May 2007.

[11] Z. Lacroix, H. Murthy, F. Naumann, and L. Raschid, "Links and paths through life sciences data sources," in *Proc. 1st Int.Workshop Data Integr. Life Sci.*, Leipzig, Germany, 2004, pp. 203–211

[12] J. Noessner,M.Niepert, C. Meilicke, andH. Stuckenschmidt, "Leveragingterminological structure for object reconciliation," in *The Semantic Web: Research and Applications*. Berlin, Germany: Springer, 2010, pp. 334–348

[13] M. E. J. Newman, "Detecting community structure in networks," Eur.Phys. J., vol. 38, pp. 321–330, Mar. 2004.

[14] J. Huan and J. Prins, "Efficient mining of frequent subgraphs in the presence of isomorphism," in Proc. 3rd IEEE Int. Conf. Data Mining, Melbourne, FL, USA, 2003, pp. 549–552

[15] D. Hand, "Principles of data mining," Drug Safety, vol. 30, pp. 621–622,Jul. 2007.

[16] J. Hans and M. Kamber, Data Mining: Concepts and Techniques, 2nd ed. Burlington, MA, USA: Morgan Kaufmann, 2006

[17] C. Corley, D. Cook, A. Mikler, and K. Singh, "Text and structural data mining of influenza mentions in web and social media," Int. J. Environ.Res. Public Health, vol. 7, pp. 596–615, Feb. 2010.

[18] S. R. Das and M. Y. Chen, "Yahoo! for amazon: Sentiment extraction from small talk on the Web," Manage. Sci., vol. 53, pp. 1375–1388, Sep. 2007.

[19] E. Riloff, "Little words can make a big difference for text classification,"in Proc. 18th Annu. Int. ACM SIGIR Conf. Res. Develop. Inf. Retrieval,Seattle, WA, USA, 1995, pp. 130–136.

[20] W. Yih, P. H. Chang, and W. Kim, "Mining online deal forums for hot deals," in Proc. IEEE/WIC/ACM Int. Conf. Web Intell., Beijing, China, 2004, pp. 384–390.

[21] B. Pang, L. Lee, and S. Vaithyanathan, "Thumbs up? sentiment classification using machine learning techniques," in Proc. Conf. Empirical Methods Natural Language Process., Philadelphia, PA, USA, 2002,pp. 79–86

[22] X. Feng *et al.*, "Assessing pancreatic cancer risk associated with dipeptidyl peptidase 4 inhibitors: Data mining of FDA adverse event reportingsystem (FAERS)," *J. Pharmacovigilance*, vol. 1, Jul. 2013.

[23] K. Y. Chan, C. K. Kwong, and T. C. Wong, "Modeling customer satisfaction for product development using genetic programming," *J. Eng. Des.*,vol. 22, no. 1, pp. 56–68, Jan. 2011.

[24] I. Frommholz and M. Lechtenfeld, "Determining the polarity of postings for discussion search," in *Proc. Workshop Woche: Lernen, WissenAdaptivit¨at*, W¨urzburg, Germany, 2008, pp. 49–56

[25] J. Schectman. (2013, May 1). Glaxo Mined Online Parent Discussion Boards For Vaccine Worries. [Online]. Available: http://blogs.wsj.com/cio/2013/05/01/glaxo-mined-online-parent-discussion-boards-forvaccine-worries.

[26] R. McBride. (2012, Aug. 1). Merck to Draw on Social Network for Psoriasis Patients. [Online]. Available: http://www.fiercebiotechit.com/story/merck-draw-social-network-psoriasis-patients/2012–08–13

[27] Five ways a Boston Children's Hospital spin-off s using social media for public health. [Online]. Available: http://mobihealthnews.com/42072/five-ways-a-boston-childrens-hospital-spin-off-is-usingsocial-media-for-public-health

[28] GSK. Merck use social media to learn how patients use drugs outside the lab. [Online]. Available: http://mobihealthnews.com/47284/gskmerck-use-social-media-to-learn-how-patients-use-drugs-outside-the-lab

[29] W. Yih and V. Qazvinian, "Measuring word relatedness using heterogenous vector space models," presented at the *Conf. North Amer. ChapterAssociation Computational Linguistics: Human Language Technologies,*Montreal, Canada, 2012.

[30] A. Sarker *et al.*, "Utilizing social media data for pharmacolvigilance: A review," *J. Biomed. Informat.*, vol. 54, pp. 202–212, 2015

[31] J. L. Wu, L. C. Yu, and P. C. Chang, "Detecting causality from online psychiatric texts using inter-sentential language patterns," *BMC Med. Informat.Decision Making*, vol. 12, 2012.

[32] L. C. Yu, C. L. Chan, C. C. Lin, and I. C. Lin, "Mining association language patterns using a distributional semantic

model for negative life event classification," *J. Biomed. Informat.*, vol. 44, pp. 509–18, 2011.

[33] M. Dumontier and N. Villanueva-Rosales, "Towards pharmacogenomics knowledge discovery with the semantic web," *Briefings Bioinformat.*,vol. 2, pp. 153–63, 2009.

[34] F. Moradi, A. M. Eklund, D. Kokkinakis, T. Olovsson, and P. Tsifas, "A graph-based analysis of medical queries of a swedish health care portal,"in *Proc. 5th Int. Workshop Health Text Mining Inf. Analysis*, Gothenburg,Sweden, Apr. 26–30, 2014, pp. 2–10.

[35] I. S. Bedmar, R. Revert, and P. Martinez, "Detecting drugs and adverse events from Spanish health social media streams," in *Proc. 5th Int. Workshop Health Text Mining Inf. Analysis*, Gothenburg, Sweden, Apr. 26–30,2014, pp. 106–115.

[36] M. Paul and M. Dredze, "Discovering health topics in social media using topic models," PLOS One, vol. 9, pp. e103–e408, 2014.

[37] K. Zhao et al., "Finding influential users of online health communities: A new metric based on sentiment influence," J. Amer. Med. Inform. Assoc.,vol. 21, pp. 212–218, 2014.

[38] K. Zhao et al., "Identifying leaders in an online cancer survivor community,"presented at the 21st Annu. Workshop Information Technologies Systems, Shanghai, China, 2011.

[39] M. De Choudhury, S. Counts, and E. Horvitz, "Social media as a measurement tool of depression in populations," in Proc. 5th Annu. ACM Web Sci. Conf., 2013, pp. 47–56.

[40] M. De Choudhury, M. Gamon, S. Counts, and E. Horvitz, "Predicting depression via social media," presented at the Int. Conf. Weblogs Social Media , Cambridge , MA, USA, 2013.

[41] A. Akay, A. Dragomir, and B. Erlandsson, "Network-based modelling and intelligent data mining of social media for improving care," IEEE J. Biomed. Health Informat., vol. 19, no. 1, pp. 210–218, Jan. 2015.

[42] I. Mierswa, M. Wurst, W. Michael, R. Klinkenberg, M. Scholz, and T. Euler, "YALE: Rapid prototyping for complex data mining tasks,"in Proc. 12th ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining,Philadelphia, PA, USA, 2006, pp. 935–940

[43] S. Bird, E.Klein, and E. Loper, Natural Language ProcessingWith Python,Sebastopol, CA, USA: O'Reilly Media, 2009, pp. 504

[44] P. Soucy and G.W. Mineau, "Beyond TFIDF scoreing for text categorization in the vector space model," in Proc.19th Int. Joint Conf. Artif. Intell.,Edinburgh, U.K., 2005, pp. 1130–1135.

[45] [Online]. Available: https://gate.ac.uk/

[46][Online]Available:http://www.nlm.nih.gov/research/uml s/

[47] Diagnostic and StatisticalManual of Mental Disorders, 5th ed., American Psychiatric Association, Washington, DC, USA, 2013.

[48] T. Kanungo et al., "An efficient k-means clustering algorithm: Analysis and implementation," Pattern Anal. Mach. Intell. Trans., vol. 24, no. 7,pp. 881–892, 2002.

[49] I. A. Maraziotis, K. Dimitrakopoulou, and A. Bezerianos, "Growing functional modules from a seed protein via integration of protein interaction and gene expression data," BMC Bioinformat., vol. 8, no. 1, p. 408, 2007.

[50] E. L. Martelot and C. Hankin, "Multi-scale community detection using stability as optimisation criterion in a greedy algorithm," in Proc. Int. Conf. Knowl. Discovery Inf. Retrieval, Paris, France, Oct. 2011, pp. 216–225.

[51] S. Fortunato, "Community detection in graphs," Phys. Rep., vol. 486,no. 3, pp. 75–174, 2010.

[52] R. Lambiotte, "Multi-scale modularity in complex networks," in Proc.IEEE Int. Symp. Modeling Optimization Mobile, Ad Hoc Wireless Netw.,2010, pp. 546–553.

[53] ibid.

[54] L. C. Tsoi et al., "Text-mining approach to evaluate terms for ontology development," J. Biomed. Informat., vol. 42, no. 5, pp. 824–830, 2009.

[55] J. M. Kleinberg, "Authoritative sources in a hyperlinked environment,"J. ACM, vol. 46, no. 5, pp. 604–632, 1999.

[56] J. M. Pujol, R. Sang¨uesa, and J. Delgado, "Extracting reputation in multi agent systems bymeans of social network topology," in Proc. 1st Int. Joint Conf. Auton. Agents Multiagent Syst., 2002, pp. 467–474.

[57] ibid

[58] D. L. Davies and D. W. Bouldin, "A cluster separation measure," IEEE
Trans. Pattern Anal. Mach. Intell., vol. PAMI-1, no. 2, pp. 224–227,
Apr. 1979.