Ontology Based Text Mining Framework for Vulnerability Assessment in Health and Social Care

Jamilah Nasution, Phong Thanh Nguyen, Muhamad Rusliyadi, E. Laxmi Lydia, K. Shankar

Abstract*: *The measure of advanced data that is made and utilized is gradually growing alongside the development of refined equipment and programming. What's more, true information arrive in a decent variety of structures and can be massively cumbersome. This has increased the requirement for great procedures that can find and uncover engaging realities and valuable data from these information. Text Mining (TM), which is an intricate procedure; has been effectively utilized for this reason. Text mining on the other hand alluded to as content data mining, pretty much identical to content examination, can be characterized as the way toward extricating great data from content. Text mining includes the way toward organizing the information, determining designs inside the organized information and in conclusion translation and disclosure of the yield. This paper gives overview on Text mining. Toward the end, this paper presents ontology system to adapt up to extreme social media textual data. We depict vulnerability assessment model intended to help patient management in health and social care. Such a framework isn't intended to supplant existing health and social assessment models but instead to supplement them by giving an all-encompassing image of the vulnerabilities looked by a given patient. Actually, it ought to be viewed as a screening tools for health and social care workers. One key part of the demonstrating structure is the capacity to give customized at this point multi-dimensional assessments of risk of dependent on fragmented data about the patient status, similar to the case in screening situations.
Keywords: Ontology, Information Retrieval, Text mining etc.

I. INTRODUCTION

Text mining, some of the time denoted to as text data mining, about equal to text analytics, refers to the way toward removing excellent data from text. Text mining as a rule includes the way toward organizing the input text inferring designs inside the organized information and to end with translation and thought of the yield. The word 'high - quality' in text mining by and large alludes to a mix of uniqueness, essentialness and intriguing quality. Text mining is huge zone in contrast with data recovery. Text mining forms as a rule

incorporate report characterization, grouping the record, building metaphysics, assessment investigation, rundown, Information extraction and so on. On the opposite side, IR manages creeping, parsing and ordering and recovering records. With an iterative methodology, an association can utilize content examination to grow knowledge into the substance explicit qualities, for example, viewpoint, quality and criticalness of the information. Text mining can be characterized as the way toward getting data from text sources.

The amount of textual data is consistently developing. Regular data mining strategies are not adequate to assess unstructured data. So there is have to utilize some semantic methodologies. Text mining forms the unstructured information determines important data. What's more, along these lines, makes the data available to the different information mining methods, including statistical and AI.

II. ONTOLOGY FRAMEWORK

Figure demonstrates the diagrammatic perspective on our framework. There are two interfaces accessible in the framework. One for the client and another for the developer. Various modules of the structure are clarified as underneath:

- Database: All the information and data are taken care of in the database. The set away information is used to answer the user's question.
- Browser: Web Browser acts an interface among customer and Programmer. Customer can use OWL structure for addressing and designer can use SPARQL.
- OWL Ontology: In this part, rules are described and made. These principles portray unmistakable cosmology classes and interrelationships between them.
- Management Services: Different standards describe different classes and through classes we give administrations. Different sorts of administrations are open like Billing, Data Usage, and Authentication, etc.
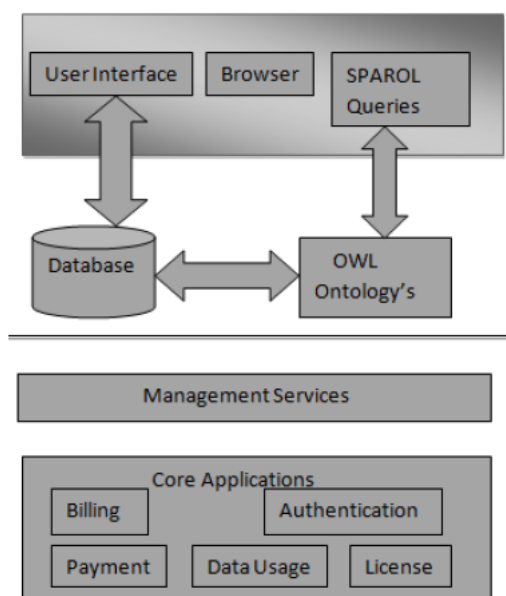
Fig1. Ontology Framework

III. ONTOLOGICAL TECHNIQUES IN HEALTH

Ontological methods have been extensively applied to medicinal and natural research. All of the seven structures examined here has been made upon some bio-ontologies. We rapidly present six such systems, trailed by the last one where more portrayal of regular point of view is given to empower the peruser to all the more probable perceive how biomedical and natural informatics may urge zone authorities to increment organic bits of learning.

Cantor et al have discussed the issue of mapping thoughts in GO to the Unified Medical Language System (UMLS). Such a mapping may allow the abuse of the UMLS semantic framework to interface various characteristics, through their remark in GO, to wonderful clinical outcomes, possibly uncovering natural associations. This assessment reveals the inborn difficulties in the joining of vocabularies made in different propensities and by experts in different fields, similarly as the characteristics of different methodology used to accomplish this coordination.

Sulman et al have uncovered a significant standards composed guide of the zone created (CompView) to recognize all markers in the most diminutive area of covering deletion. A commonplace physical cell cross breed board is used to constrain even more completely those markers recognized in CompView as inside or covering the region, and a gathering from clones is used to favor STS content by electronic PCR and to perceive transcripts. The makers have assumed that the clarification of a putative tumor silencer locus gives an advantage for support examination of meningioma competitor qualities.

Birkland et al have shown a structure, Biozon, to address the issues experienced in the compromise of heterogeneous data types in the science space. Biozon offers researcher another knowledge resource to explore through and investigate by binding together different organic databases that comprise of an assortment of information types (e.g., DNA groupings, proteins, associations, and cell pathways). Biozon is not quite the same as past endeavors as in it utilizes a solitary broad and firmly associated diagram mapping wrapped with progressive philosophy of archives and relations. Past warehousing

existing information, Biozon processes and stores novel inferred information, likeness connections and utilitarian forecasts, for instance. The combination of comparability information permits proliferation of knowledge by inference and fuzzy searches.
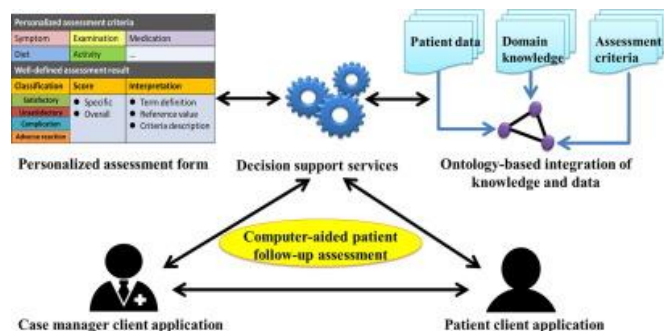

Fig 2: Ontological Techniques in Health

The estimation of any sort of information is enormously improved when information exist in a structure permitting the joining with other information. One way to deal with mix is through the comment of different collections of information utilizing basic controlled vocabularies or ontologies. Tragically, the very achievement of this methodology has prompted an expansion of ontologies, which itself makes deterrents to coordination. So as to beat such issues, Smith et al have depicted a technique, specifically, the OBO Foundry activity. The long haul objective of the OBO activity is that the information produced through biomedical research should frame a solitary, steady, in total extending, and algorithmically tractable entirety. Endeavors to understand this objective are still particularly in the demonstrating stage. By the by, the underlying endeavors mirror an endeavor to push it between the adaptability that is imperative to logical development and the organization of rules that is essential to fruitful coordination.

Huang et al have displayed a space explicit learning base based upon the Ontology for MicroRNA Targets (OMIT) to encourage information securing in the field of miRNA target quality expectation. The recognizable proof and portrayal of significant jobs that miRNAs perform in human disease has progressively turned into a functioning examination zone. Be that as it may, the forecast of miRNA target qualities stays a provoking errand to malignancy specialists. Current expectation procedures are tedious, blunder inclined, and subject to researcher's constrained earlier learning. The OMIT framework intends to help researcher in unwinding significant jobs of miRNAs in human disease; subsequently, OMIT can enable clinicians to settle on steady choices when treating malignant growth patients. To be progressively explicit, it is outstanding that each miRNA can have several conceivable objective qualities. At present, there are various objective expectation databases that are geologically disseminated worldwide and that have embraced very various constructions and wordings. In addition, by and large, extra data for objective qualities is basic for scientists to see completely these qualities' natural capacities. As a general rule, such extra data isn't accessible in objective expectation databases. Rather, different assets, for example, the GO ontologies are required for this reason.

Taking mir-21 for instance, miRDB, TargetScan, and PicTar report 245, 211, and 176 objective qualities for mir-21, separately. It is exceptionally testing, if certainly feasible, for scientists to look through physically a sum of 733 competitor target qualities, let alone to further scan for helpful data on every quality covered up in GO. Truth be told, the circumstance could be surprisingly more terrible: researcher as a rule utilize multiple databases in the miRNA inquire about territory. To deal with this test, the OMIT structure enables scholars to find miRNA up-and-comer target qualities in a substantially more effective way: (1) learning from different databases is naturally acquired, incorporated, and exhibited to clients; and (2) related data from GO is accommodated each recovered objective quality. Thusly, scientists can spare a lot of time that would have been spent if a manual inquiry were to be done. Furthermore, because of derivation motors (otherwise called metaphysics reasoners) explicitly intended for OWL ontologies, OMIT can distinguish shrouded learning that isn't express in the first information. Correspondingly, another end, "mir-21 advances Tumor", can be promptly acquired also. These additional ends will assist scientists with generalizing their discoveries to progressively demonstrate frameworks.

IV. TEXT MINING IN CLINICAL AND BIOMEDICAL DATA

Semantic decision genuinely strong systems can upgrade semantic data bases as discretionary sources to help specialists in choosing sound clinical decisions. Lin et al [13] have investigated a clinical confirmation recuperation structure and have guessed that social event recouped MEDLINE references into semantically clear packs, in perspective on normally removed intercessions from the calculated substance, addresses a practical framework for showing results, differentiated and a standard situated rundown. In light of this hypothesis, they have arranged a work procedure that: (1) perceives all components having a spot with synthetics and meds, contraptions, and methods from recuperated altered works; (2) removes major interventions by consigning each intercession (and the related hypothetical) to its own one of a kind gathering; (3) iteratively unions bundles whose interventions share a run of the mill UMLS hypernym, rising the UMLS hierarchy of leadership at the same time; and (4) sorts realizes the solicitation for the first PubMed results inside each formed gathering.
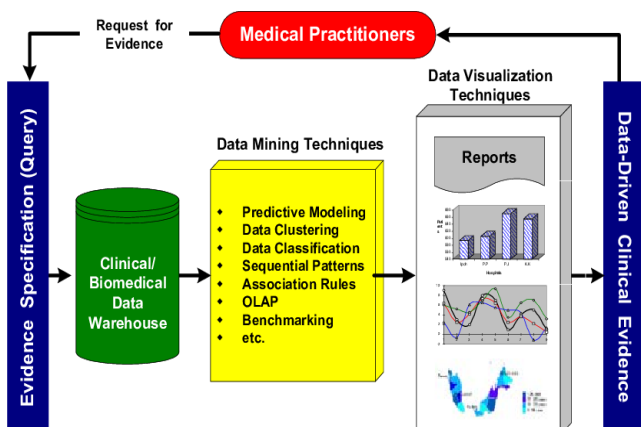
V. VULNERABILITY ASSESSMENT IN HEALTH AND SOCIAL CARE

As this health and therapeutic information innovation framework rises and advances, a large group of dangers to security and protection will in like manner rise and develop. Social insurance and medicinal frameworks are an extremely unpredictable interchange of advancements, individuals, approach, and enactment. Notwithstanding, these frameworks work in a particularly testing security condition since social insurance and restorative information must be caught and recovered at the purpose of-care, that is, it must be portable and remote.

This segment centers on dangers to security and protection of human services and medicinal information regarding security and security dangers. A danger can be characterized as the misuse of framework helplessness. Helplessness is characterized as "an imperfection or shortcoming in framework security methodology, structure, execution, or inner controls that could be abused to achieve a security break or an infringement of the framework's security strategy".

Landry et al have built up a danger tree to survey and oversee dangers to medicinal services and therapeutic information. Kotz has proposed a system that sorts out a lot of 25 dangers by character dangers, get to dangers, and divulgence dangers. Samy et al have recognized 22 classifications of health data frameworks dangers. Their exploration has recognized five basic territories, specifically, control disappointment/misfortune, demonstrations of human mistake or disappointment, mechanical oldness, equipment disappointments or blunders, and programming disappointments or blunders. The focal point of this writing survey is vulnerabilities related with unapproved control, information misfortune, and information defilement. Results of misusing security vulnerabilities incorporate introduction to monetary damage, mental anguish, social disgrace, wholesale fraud, and poor healthcare and medical results.

Digital protest is an outflow of a social development using data advances. Digital dissidents have since quite a while ago focused on dubious medicinal services, for example, abortion or medical animal research.

In the event that improperly actualized, data innovation can really prompt healthcare and therapeutic blunders that speak to a unintended misuse of vulnerability. For instance, late examinations have demonstrated that PC doctor request passage (CPOE) frameworks can encourage prescription and request entry error. There are two noteworthy foundations for this kind of mistake. Initially, healthcare and medicinal information originate from a wide scope of sources and in fluctuating organizations. In the event that medicinal information are not effectively incorporated in the database or improperly compared on a PC screen, restorative blunders can happen. Second, the presentation of data innovation into the progression of purpose of-care forms definitely changes and influences the ordinary progression of occasions. Guardians work in a performed multiple tasks, time-sensitive, and frantic condition.

In the event that the creation and recovery of social insurance and therapeutic information ineffectively coordinates the typical progression of medicinal services conveyance, time-delicate treatments, for example, early revival might be postponed to the disadvantage of patients' health and chances of survival. Creating "carefully designed" components for the data innovation that oversees human services and medicinal information is a nontrivial challenge. Therapeutic devices must be associated with different devices and servers through remote correspondences, accordingly, the information transmitted by means of remote systems are defenseless against both security and protection dangers.

Next-generation information technology framework should build the accessibility and exactness of healthcare and medical data. Be that as it may, unapproved access is as yet a genuine powerlessness. There are numerous types of and explanations behind unapproved access, for example, gaining medicinal treatment with an alternate individual's protection approach. Such unapproved access is viewed as a type of data fraud. Note that a typical helplessness emerges from the very idea of social insurance conveyance.

VI. CONCLUSION

Innovative computing procedures based upon the inexorably unavoidable digital foundation are required all together for medicinal researchers and scholars to get an improved capacity to incorporate, share, and reuse initially heterogeneous information from disseminated labs. The proportion of cutting edge information that is made and used is step by step developing nearby the improvement of refined hardware and programming. Furthermore, genuine data land in a respectable assortment of structures and can be enormously unwieldy. This has expanded the necessity for extraordinary methods that can discover and reveal drawing in substances and significant information from these data. Content Mining (TM), which is a complex system; has been adequately used thus. Content mining then again suggested as substance information mining, basically indistinguishable from substance assessment, can be described as the route toward removing extraordinary information from substance. Content mining incorporates the path toward sorting out the data, deciding plans inside the composed data and in end interpretation and exposure of the yield. This paper gives diagram on Text mining. Close to the end, this paper presents philosophy framework to adjust up to extraordinary online networking printed information. We portray weakness appraisal model expected to help tolerant administration in wellbeing and social consideration. Such a system isn't proposed to displace existing wellbeing and social appraisal models however rather to enhance them by giving a widely inclusive picture of the vulnerabilities looked by a given patient. In reality, it should be seen as a screening apparatuses for wellbeing and social consideration laborers. One key piece of the showing structure is the ability to give tweaked now multi-dimensional appraisals of danger of subject to divided information about the patient status, like the case in screening circumstances. The objective of this paper is to give perusers a prologue to significant computing subjects that can enable researchers to get a superior understand the value of biological functions at various standards.

REFERENCES

1. Nguyen, "Text mining and Network Analysis of Digital Libraries in Data Mining Applications with R", (2014).
2. He W., Zha S. and Li L., "Social media competitive analysis and text mining", (2013).
3. Adam Bosworth. Putting health into the patient's hands - consumerism and health
4. care. Opening Plenary Session and Keynote Address, May 2007.
5. S. H. Brenner and E. J. Mckinin. CINAHL and MEDLINE: a comparison of indexingpractices. Bulletin of the Medical Library Association, 77(4):366{371, October 1989.
6. MagalyDouyere, Lina F. Soualmia, AurelieNeveol, Alexandrina Rogozan, BadisseDahamna, Jean-Philippe Leroy, Benoit Thirion, and Stefan J. Darmoni. Enhancingthe MeSH thesaurus to retrieve French online health resources in a quality-controlledgateway. Health Information and Libraries Journal, 21(4):253{261, 2004.
7. Haseeb, M., Kot, S., Hussain, H., & Jermsittiparsert, K. 2019. "Impact of Economic Growth, Environmental Pollution, and Energy Consumption on Health Expenditure and R and D Expenditure of ASEAN Countries." Energies 12 (19): 3598.
8. Huda, S., Tsani, I., Syazali, M., Umam, R., & Jermsittiparsert, K. 2020. "The Management of Educational System Using Three Law Auguste Comte: A Case of Islamic Schools." Management Science Letters 10 (3) (In press), DOI: 10.5267/j.msl.2019.9.018.
9. Usak, M., Kubiatko, M., Shabbir, M., Dudnik, O., Jermsittiparsert, K., & Rajabion, L. 2019. "Health Care Service Delivery Based on the Internet of Things: A Systematic and Comprehensive Study." International Journal of Communication Systems 32 (14): e4179.
10. Jermsittiparsert, K., Ambarita, D., Mihardjo, L., & Ghani, E. 2019. "Risk-Return through Financial Ratios as Determinants of Stock Price: A Study from ASEAN Region." Journal of Security and Sustainability Issues 9 (1): 199-210.
11. Thabhiranrak, T. & Jermsittiparsert, K. 2019. "Towards Sustainable Functioning of Organization: Women Empowerment and Corporate Management Culture." Journal of Security and Sustainability Issues 9 (1): 321-332.
12. Chienwattanasook, K. & Jermsittiparsert, K. 2019. "Impact of Entrepreneur Education on Entrepreneurial Self-Employment: A Case Study from Thailand." Polish Journal of Management Studies 19 (1): 106-116.
13. Jermsittiparsert, K., Sutduean, J., Sriyakul, T., & Khumboon, R. 2019. "The Role of Customer Responsiveness in Improving the External Performance of an Agile Supply Chain." Polish Journal of Management Studies 19 (2): 206-217.
14. Jermsittiparsert, K., Sutduean, J., & Sriyakul, T. 2019. "Effect of Service Innovation and Market Intelligence on Supply Chain Performance in Indonesian Fishing Industry." Industrial Engineering & Management Systems 18 (3): 408-417.
15. Jermsittiparsert, K., Namdej, P., & Somjai, S. 2019. "Green Supply Chain Practices and Sustainable Performance: Moderating Role of Total Quality Management Practices in Electronic Industry of Thailand." International Journal of Supply Chain Management 8 (3): 33-46.
16. Somjai, S. & Jermsittiparsert, K. 2019. "The Trade-off between Cost and Environmental Performance in the Presence of Sustainable Supply Chain." International Journal of Supply Chain Management 8 (4): 237-247.
17. Jermsittiparsert, K. & Sawasdee, A. 2012. "Formal Education for Non-Thai or Undocumented Person in Thailand amidst the Challenge of Nationalism and Transnationalism: A Case Study of Wat Sirimongkhol School, Samut Sakhon Province." Kasetsart Journal - Social Sciences 33 (2): 203-213.
18. Chienwattanasook, K., Wattanapongphasuk, W., Prianto, A., & Jermsittiparsert, K. 2019. "Corporate Entrepreneurship and Business Performance of Logistic Companies in Indonesia." Industrial Engineering & Management Systems 18 (3): 538-547.
19. Dawabsheh, M., Hussein, A., & Jermsittiparsert, K. 2019. "The Triangular Relationship between TQM, Organizational Excellence and Organizational Performance: A Case of Arab American University Palestine." Management Science Letters 9 (6): 921-932.

20. Jermsittiparsert, K., Siam, M., Issa, M., Ahmed, U., & Pahi, M. 2019. "Do Consumers Expect Companies to Be Socially Responsible? The Impact of Corporate Social Responsibility on Buying Behavior." Uncertain Supply Chain Management 7 (4): 741-752.

21. Syazali, M., Putra, F., Rinaldi, A., Utami, L., Widayanti, Umam, R., & Jermsittiparsert, K. 2019. "Partial Correlation Analysis Using Multiple Linear Regression: Impact on Business Environment of Digital Marketing Interest in the Era of Industrial Revolution 4.0." Management Science Letters 9 (11): 1875-1886.

22. Sae-Lim, P. & Jermsittiparsert, K. 2019. "Is the Fourth Industrial Revolution a Panacea? Risks toward the Fourth Industrial Revolution: Evidence in the Thai Economy." International Journal of Innovation, Creativity and Change 5 (2): 732-752.

23. Chatchawanchanchanakij, P., Arpornpisal, C., & Jermsittiparsert, K. 2019. "The Role of Corporate Governance in Creating a Capable Supply Chain: A Case of Indonesian Tin Industry." International Journal of Supply Chain Management 8 (3): 854-864.

24. Hartinah, S., Suharso, P., Umam, R., Syazali, M., Lestari, B., Roslina, R., & Jermsittiparsert, K. 2020. "Teacher's Performance Management: The Role of Principal's Leadership, Work Environment and Motivation in Tegal City, Indonesia." Management Science Letters 10 (1): 235-246.

25. Haseeb, M., Hussain, H., Slusarczyk, B., & Jermsittiparsert, K. 2019. "Industry 4.0: A Solution towards Technology Challenges of Sustainable Business Performance." Social Sciences 8 (5): 184.

26. Haseeb, M., Hussain, H., Kot, S., Androniceanu, A., & Jermsittiparsert, K. 2019. "Role of Social and Technological Challenges in Achieving a Sustainable Competitive Advantage and Sustainable Business Performance." Sustainability 11 (14): 3811.

27. Tran Dung and WataruKameyama. Ontology-based information extraction and informationretrieval in health care domain. Data Warehousing and Knowledge Discovery,pages 323{333, 2007.

28. William R. Hersh, Henning Muller, Je_ery R. Jensen, Jianji Yang, Paul N. Gorman,and Patrick Ruch. Advancing biomedical image retrieval: Development and analysisof a test collection. J Am Med Inform Assoc, 13(5):488{496, September 2006.

29. AngelosHliaoutakis, GiannisVarelas, Euripides Petrakis, and EvangelosMilios. Med-Search: A retrieval system for medical information based on semantic similarity. Re-search and Advanced Technology for Digital Libraries, 4172/2006:512{515, 2006.