

{tag} International Journal of Computer Applications
Foundation of Computer Science (FCS), NY, USA

[Volume 179](#)

-
[Number 17](#)

Year of Publication: 2018

Authors:

Ahmed Y. Abu El-Qumsan, Alaa M. Elhalees

10.5120/ijca2018916301

{bibtex}2018916301.bib{/bibtex}

Abstract

The torrential information in the medical records is considered a great problem because it difficult to distinguish the needed and necessary information from the huge quantity of data. As a result, the importance of summarize medical reports is growing day after day. Medical information extraction is one of the important topics that aim to identify medical information and detect hidden relations. This topic is considered one of the most important topics in the field of text mining where is used to process unstructured texts and extract meaningful information which is hidden in the unstructured texts.

The information extracted from medical reports is very useful to medical staff to detect hidden relations between medical information, and making decisions that will improve the medical service for patients, in addition to saving time and effort.

In our paper, an approach that use template based medical reports summarization has been developed to transfer medical reports from semi structured and unstructured form to structured

form. It classifies the identified entities then extracts important information such as diseases, medical procedures, and drugs. After that, it can discover hidden relationships between medical information by using association rules. The dataset used in this paper was collected from the Palestinian Ministry of Health.

To evaluate the performance and effectiveness of our model, human expert has been used as a reference to measure the degree of acceptance of the extracted association rules which have been extracted from the dataset. So, Likert's scale has been used for evaluation. After the data analysis obtained from the questionnaire. It shows us that the proportion of accuracy association rules, which have been extracted is about 80%.

References

1. Iahari, E., Kumar, D., and Ubale, M. 2014. A Comprehensive Survey on Feature Extraction in Text Summarization. *Int.j. computer Technology & Application*, Vol 5 (1), 248-256.
2. Radev, D. and Mckeown, K. 2002. Introduction to the special issue on summarization. *Computational Linguistics*, 28 (4), 339 – 408.
3. Mani, I., and Maybury, M. 2001. *Advances in Automatic Text Summarization*. The MIT Cambridge, Massachusetts London, England.
4. Bunescu, M., Ge R., Mooney, R., Marcotte, E., and Ramani, A. 2002. *Extracting Gene and Protein Names from Biomedical Abstracts*. Unpublished Technical Note.
5. Jung, J., and Jo, G. 2003. Template-Based E-mail Summarization for Wireless Devices. *computer and information sciences - ISCIS, LNCS 2869*, 99–106.
6. Deléger, L., Grouin, C., & Zweigenbaum, P. 2010. Extracting medical information from narrative patient records: the case of medication-related information. *J Am Med Inform Assoc*, 17(5): 555–558.
7. Jonnagaddala, J., Liaw, S., Ray, P., Kumar, M., Chang, N., and Dai, H. 2015. Coronary artery disease risk assessment from unstructured electronic health records using text mining. *Journal of Biomedical Informatics* 58, S203–S210.
8. Chen, A., and Verma, R. 2006. A query-based medical Information summarization system Using Ontology Knowledge. In the proceedings of the 19th IEEE Symposium on Computer based Medical Systems.
9. Sarkar, K. 2009. Using Domain Knowledge for Text Summarization in Medical Domain. *International Journal of Recent Trends in Engineering*, 1 (1).
10. Xu, H. Stenner, S., Doan, S., Johnson, K., Waitman, L., & Denny, J. 2010. MedEx: a medication information extraction system for clinical narratives. *Journal of the American Medical of Informatics Associations*, 17(1):19-24.
11. Gold, S., Elhadad, N., Zhu, X., Cimino, J., & Hripcsak, G. 2008. Extracting Structured Medication Event Information from Discharge Summaries. *AMIA Annual Symposium Proceeding Archive.*, 237–241.
12. Doddi, S., Marathe, A., Ravi, S., & Torney, D. 2001. Discovery of Association Rules in Medical Data. *Med. Inform. Internet. Med.*, 26, 25–33.
13. Rashid, M., Hoque, M., & Sattar, S. 2014. Association Rules Mining Based Clinical Observations. *Bioinformatics*, 9(11): 555–559.
14. Ordonez, C., Omiecinski, E., Braal, L., Santana, C., Ezquerro, N., Taboada, J.,

rawczynska, E. 2001. Mining Constrained Association Rules to Predict Heart Disease. Proceeding ICDM '01 Proceedings of the 2001 IEEE International Conference on Data Mining, 433-440.

15. Nahar, J., Imam, T., Tickle, K., & Chen, Y. 2013. Association rule mining to detect factors which contribute to heart disease in males and females. *Expert Systems with Applications* 40, 1086–1093.

16. Greenwood, M., Roberts, A., Aswani, N., & Gooch, P. 2012. Initial prototype for semantic annotation of the Khresmoi literature. project deliverable Khresmoi.

17. Spasić, I., Zhao, B., Jones, C., & Button, K. 2015. KneeTex: an ontology-driven system for information extraction from MRI reports. *Journal of Biomedical Semantics*.

18. Sproat, R., Black, A., Chen, S., Kumar, S., Ostendorf, M., & Richards, C. 2001. Normalization of non-standard words. *Computer Speech and Language* 15, 287–333.

19. Elsebai, A. 2009. A rules based system for named entity recognition in modern standard. University of Salford.

20. Mukund, S., Srihari, R., & Peterson, E. 2010. An Information-Extraction System for Urdu—A Resource-Poor Language. *ACM Transactions on Asian Language Information Processing*, 9 (4).

21. Al-Zaidy, R., Fung, B., Youssef, A., & Fortin, F. 2012. Mining criminal networks from unstructured text documents. *Digital Investigation*, 8 (3-4), 147-160.

22. Jiang, J. 2012. Information extraction from text. C.C. Aggarwal, C. Zhai (Eds.), *Mining text data*, Springer, 11–41.

23. Doddi, S., Marathe, A., Ravi, S., & Torney, D. 2001. Discovery of Association Rules in Medical Data. *Med. Inform. Internet. Med.*, 26, 25–33.

24. Ozen, G., Yaman, M., & Acar, G. 2012. Determination of the employment status of graduates of recreation department. *The Online Journal of Recreation and Sport*, 1 (2).

Index Terms

Computer Science

Biomedical

Keywords

Template Based Summarization, Text Mining, Information Extraction, Medical Reports, Named Entity Recognition, Association Rules.