Abstract

In this paper, we are interested in cohort search, representation, and prediction. Information retrieval and text mining technique were proposed based on Term Frequency Inverse Document Frequency (TF.IDF) to extract important terms. Also, a formal and algorithmic model was formulated to compute: readable, concise cohorts of patients and find similarities between patient trajectories.

Finally, Patient health trajectories were analyzed using a Deep Learning architecture from intensive experimental processes based on two parallel Minimal Gated Recurrent Unit networks, working in a bi-directional manner. The obtained result shows an improvement in the performance of computer-aided medicine and serves as a guide in designing artificial neural networks used in prediction tasks.

References
1. E. Balka, “Electronic patient records,” Nurs. Manage., vol. 5, no. 4, pp. 39–39, 2016.
2. J. G. Beun, “Electronic healthcare record; a way to empower the patient,” Int. J. Med. Inform., vol. 69, no. 2–3, p. 191—196, Mar. 2003.
3. B. Omidvar-Tehrani, S. Amer-Yahia, and L. V. S. Lakshmanan, “Cohort representation and exploration,” Proc. - 2018 IEEE 5th Int. Conf. Data Sci. Adv. Anal. DSAA 2018, pp. 169–178, 2019.
4. T. Pham, T. Tran, D. Phung, and S. Venkatesh, “Predicting healthcare trajectories from medical records: A deep learning approach,” J. Biomed. Inform., vol. 69, pp. 218–229, 2017.
5. S. Rodrigues-Jr, JF; Spadon, G; Machado, BB; Amer-Yahia, “Patient trajectory prediction in the Mimic-III dataset challenges and pitfalls,” 35ème Conférence sur la Gest. Données – Principes, Technol. Appl., 2019.
6. S. Qaiser and R. Ali, “Text Mining: Use of TF-IDF to Examine the Relevance of Words to Documents Text Mining: Use of TF-IDF to Examine the Relevance of Words to Documents,” no. July 2018.
7. K. Hans, “Das System Kaliumsulfai-Kaliurnsulfid,” Zeitschrift für Anorg. und Allg. Chemie, vol. 164, pp. 45–56, 1931.
8. E. S. Ristad and P. N. Yianilos, “Learning string-edit distance,” IEEE Trans. Pattern Anal. Mach. Intell., vol. 20, no. 5, pp. 522–532, 1998.
9. A. Heuser, M. Huynh, C. Joshua, and J. C. Chang, “Asymptotic convergence in distribution of the area bounded by Kaplan-Meier curves using empirical process modeling,” 2014.
10. W. David and D. H. Rich, “Book, Software, and Web Site Reviews,” no. 11, pp. 2219–2221, 2005.
11. S. B. Needleman and C. D. Wunsch, “A general method applicable to the search for similarities in the amino acid sequence of two proteins,” J. Mol. Biol., vol. 48, no. 3, pp. 443–453, 1970.
12. X. Glorot and Y. Bengio, “Understanding the difficulty of training deep feedforward neural networks,” in Proceedings of the Thirteenth International Conference on Artificial Intelligence and Statistics, 2010, vol. 9, pp. 249–256.
13. R. Pascanu, C. Gulcehre, K. Cho, and Y. Bengio, “How to Construct Deep Recurrent Neural Networks,” pp. 1–13, 2013.
14. M. Carbin, “THE LOTTERY TICKET HYPOTHESIS:” pp. 1–42, 2019.
15. M. Schuster and K. K. Paliwal, “Bidirectional Recurrent Neural Networks,” vol. 45, no. 11, pp. 2673–2681, 1997.
16. G. A. Miller, “Human memory and the storage of information,” {IRE} Trans. Inf. Theory, vol. 2, no. 3, pp. 129–137, 1956.
17. A. Rajkomar et al., “Scalable and accurate deep learning with electronic health records,” npj Digit. Med., vol. 1, no. 1, p. 18, 2018.
18. E. Choi, M. T. Bahadori, A. Schuetz, W. F. Stewart, and J. Sun, “Doctor AI: Predicting Clinical Events via Recurrent Neural Networks,” JMLR Workshop Conf. Proc., vol. 56, pp. 301–318, Aug. 2016.

Index Terms

Computer Science

Information Sciences
Keywords

TF-IDF, EMR, Cohort, Neural Networks, Deep Learning, Patient Trajectory