Case Study #5 – Diabetes

CS6440 Intro Health Informatics, 2020 Spring

Xia, Hui
903459648
Georgia Institute of Technology

1. **Diabetes is at pandemic levels. Describe two areas of technology innovation**

    A. **Technology innovation in diabetes prevention.** Type II diabetes (T2D) represent ~95% of all the diabetes cases, and could be prevented by changing life style and/or using medicine intervention. However, T2D patients typically cannot be diagnosed 4–7 years after the onset of the disease (Harris, Klein, Welborn, & Knuiman, 1992), until noticeable vascular complication is developed. Thus, to prevent diabetes among the whole population, it is necessary to screen out the individuals with high risk of developing Type II diabetes, and those who has early stage of diabetes under development, before complication appears.

    Due to the severe social impact of T2D, a great amount of data on the features of diabetes patient is generated (Kavakiotis et al., 2017). This provide the researchers a great chance to perform machine learning (ML), especially supervised ML approaches to perform early diagnosis of T2D. Mani et. al. applied various ML algorithms, such as random forest (RF), KNN and SVM, to evaluate the diabetes risk using individual patient's electronic medical records (EMRs). Using the RF algorithm, they are able to achieve 80% Area under the Curve of ROC for predicting T2D 365 days and 180 days prior to diagnosis of diabetes using conventional medical standards (Mani, Chen, Elasy, Clayton, & Denny, 2012).

    B. **Technology innovation in diabetes treatment.** Computational technology could help in drug screening for diabetes. Recently, gene therapy drugs have been used in treatment of diabetes mellitus. The gene therapy drugs are usually inhibitors that could block certain gene pathways that are related to diabetes. For example, as Type I diabetes (T1D) is due to a loss of immune tolerance to islet antigen, there is interest in developing gene therapy drugs to re-establish it (Prud'Homme, Draghia-Akli, & Wang, 2007). Recently, computed molecular descriptors are used to find possible genetic therapy pathways for diabetes. Shoombuatong et. al. implemented a decision tree ML model, which suggests that dipeptidyl peptidase-4 (DPP4) could be a promising therapeutic route for the treatment of T2D, as it regulates glucose homeostasis (Shoombuatong et al., 2015). Similarly, a Quantitative Structure-Activity Relationship (QSAR) model was used to find the relationship between diabetes complications and activities of aldose reductase (AR), an enzyme of the polyol pathway (Patra & Chua, 2011). Currently, computational technology is still limited in finding possible gene, and their corresponding inhibitors, that are related with diabetes, chemical drugs that will actually be used in diabetes treatment are to be developed once promising gene/inhibitor pathway are suggested by these models.

2. **Explain what is novel and what will be required for these innovations to succeed.**

Among the technology that we have discussed in **Section 1**, the innovations focus on applying ML algorithms in diabetes-related data. To ensure these innovations succeed, the researchers need notice the interpretability and predictively trade-off in their ML models. It is well-known that these is a trade-off between prediction power and interpretability for machine learning models. Black box models such as SVM and deep learning could reach high prediction accuracy, but are in general not helpful in improving our understanding on 'why so?', as the reasons on why any given prediction is made is usually difficult, if not impossible, to understand. Thus, ML algorithms with good interpretability and reasonable prediction power, such as Naïve Bayes, are more suitable for diabetes diagnosis and diabetes treatment. This is because that such models can help doctors to know both the prediction and why the predictions are made, i.e. help correlate patient symptoms with test results (Kim, Cho, & Oh, 2017).

3. **How could this chronic illnesses be prevented on a pandemic level and what technologies could be used to do so?**

    A. **Computational technology can help to educate the patient on awareness of diabetes using big data.** It has been widely agreed among the healthcare realm that physical exercise is one of the best kinds of non-pharmacological treatments to prevent and control T2D (Asano et al., 2014), and that healthier lifestyle should be promoted (Alouki, Delisle, Bermúdez-Tamayo, & Johri, 2016). Thus, promoting physical exercise and healthier lifestyle become effective in preventing diabetes on a pandemic level. Currently, big data have been applied on personalized internet search engine (e.g. for advertisement purpose) (Couldry & Turow, 2014). Thus, it is rather feasible to "catch and educate" (potential and existing) diabetes patients, especially the patients that have history of in unhealthy lifestyle, by showing healthier alternatives when they are looking for certain consumptions such as food online on internet-based searching engine such as Google.

    B. **Computational tools can help to screen the patient that develop diabetes on physician visits, and inform such information to the corresponding healthcare provider.** Sharing such information to their corresponding health provider could be helpful in patient educating, as patients tend to listen and follow the suggestions made by the PCP and nurse they are familiar with, rather than following other information source (Haas, Leiser, Magill, & Sanyer, 2005).

4. **Share and describe 5 solutions for prevention, monitoring, or treating Diabetes for either type**

    A. As described in **Section 1A**, we can prevent and control the development of diabetes using machine learning algorithms to perform early diagnostics.

B. As described in **Section 1B**, we can develop drugs used in treatment of diabetes using machine learning algorithms.

C. As described in **Section 3A**, we can prevent and control the development of diabetes using big data and machine learning to feed information that promote healthier life style to diabetes patients and potential diabetes patients.

D. As described in **Section 3B**, we can prevent and control the development of diabetes using big data and machine learning to inform health care provider to helpful in educating diabetes patient and potential diabetes patients.

E. Other than relying on health care provider, as diabetes is a widely studied disease, it is possible to develop computer-aided diagnosis (CAD) systems to track individual health information, such as blood sugar level and blood pressure, which could be measured using wearable devices. The CAD system can be used to perform real-time monitoring on the risk of diabetes of the device wearer, and make appropriate health care suggestions, or even report potential risk to health care providers (Mookiah et al., 2013).

References

Alouki, K., Delisle, H., Bermúdez-Tamayo, C., & Johri, M. (2016). Lifestyle interventions to prevent type 2 diabetes: a systematic review of economic evaluation studies. *Journal of diabetes research, 2016*.

Asano, R. Y., Sales, M. M., Browne, R. A. V., Moraes, J. F. V. N., Júnior, H. J. C., Moraes, M. R., & Simões, H. G. (2014). Acute effects of physical exercise in type 2 diabetes: a review. *World journal of diabetes, 5*(5), 659.

Couldry, N., & Turow, J. (2014). Advertising, big data and the clearance of the public realm: marketers' new approaches to the content subsidy. *International Journal of Communication, 8*, 1710-1726.

Haas, L. J., Leiser, J., Magill, M. K., & Sanyer, O. N. (2005). Management of the difficult patient. *American family physician, 72*(10), 2063-2068.

Harris, M. I., Klein, R., Welborn, T. A., & Knuiman, M. W. (1992). Onset of NIDDM occurs at least 4–7 yr before clinical diagnosis. *Diabetes care, 15*(7), 815-819.

Kavakiotis, I., Tsave, O., Salifoglou, A., Maglaveras, N., Vlahavas, I., & Chouvarda, I. (2017). Machine learning and data mining methods in diabetes research. *Computational structural biotechnology journal, 15*, 104-116.

Kim, S. J., Cho, K. J., & Oh, S. (2017). Development of machine learning models for diagnosis of glaucoma. *PloS one, 12*(5).

Mani, S., Chen, Y., Elasy, T., Clayton, W., & Denny, J. (2012). *Type 2 diabetes risk forecasting from EMR data using machine learning.* Paper presented at the AMIA annual symposium proceedings.

Mookiah, M. R. K., Acharya, U. R., Chua, C. K., Lim, C. M., Ng, E., Laude, A. J. C. i. b., & medicine. (2013). Computer-aided diagnosis of diabetic retinopathy: A review. *43*(12), 2136-2155.

Patra, J. C., & Chua, B. H. (2011). Artificial neural network-based drug design for diabetes mellitus using flavonoids. *Journal of computational chemistry, 32*(4), 555-567.

Prud'Homme, G., Draghia-Akli, R., & Wang, Q. (2007). Plasmid-based gene therapy of diabetes mellitus. *Gene therapy, 14*(7), 553-564.

Shoombuatong, W., Prachayasittikul, V., Anuwongcharoen, N., Songtawee, N., Monnor, T., Prachayasittikul, S., . . . Nantasenamat, C. (2015). Navigating the chemical space of dipeptidyl peptidase-4 inhibitors. *Drug design, development therapy, 9*, 4515.