

Ali Rivzi, Alwin Baby, Kelly Holtzman, and Mansi Patel

Faculty of Engineering and Applied Science, University of Regina

ENSE 496AC: Artificial Intelligence

Dr. Kin-Choong Yow

September 25, 2020

Project Proposal

Our team proposes to create an application to support the self-diagnosing of illnesses called Medi-App; where users of the application do not need to search through clinical websites such as www.mayoclinic.org to investigate their possible mystery illness. Instead, we propose that users give their symptoms to the application to sort out. We intend to use a reasonably fast and accurate algorithm (within some example, educational allowance) to achieve legitimate results from the user's input symptoms to the suggested output illnesses.

We've begun investigating three state-of-the-art approaches for our purposes: utilizing the Bayesian network probabilistic graphical model with a test set of data representing the probability of a particular symptom relating to a particular illness; creating an Artificial Neural Network (ANN) with some naive function for the hidden layer representing the same probability; and k-means clustering where each additional user's input is a means to train the program without the need for existing legitimate medical data.

As we, the team are not medical professionals we do not have accurate and legitimate symptoms-to-illness relationship data at our disposal. We intend to retrieve legitimate data from reputable clinical websites such as www.mayoclinic.org (under the clause that we do not use the data for commercial use). We will retrieve the symptoms of illnesses using the web scraping technique aided by the Python library `lxml` and the Python module `requests` for HTML and XML handling. We also envision the use of further Python libraries and modules when implementing our algorithm(s) but have not chosen any at this time.

In terms of the Final Report, each group member's responsibilities are envisioned as follows:

1. (Kelly Holtzman) Investigation of the state-of-the-art of Bayesian networks and if the algorithm is fit for our purpose

2. (Mansi Patel) Investigation of the state-of-the-art of ANN and if the algorithm is fit for our purpose
3. (Ali Rivzi) Investigation of the state-of-the-art of k-means clustering and if the algorithm is fit for our purpose
4. (Mansi Patel) Implementation of the application component responsible for gathering legitimate symptom-to-illness data from reputable websites
5. (Alwin Baby) Implementation of the application component responsible for gathering and cleaning the user's input symptoms
6. (Ali Rivzi, Alwin Baby, Kelly Holtzman, Mansi Patel) Implementation of the algorithm deemed most fit for purpose, with responsibilities assigned as the algorithm is broken down in increments (with weight given to team members as considered fair)
7. (Ali Rivzi, Alwin Baby, Kelly Holtzman, Mansi Patel) Assessment of the implementation after testing of the program
8. (Ali Rivzi, Alwin Baby, Kelly Holtzman, Mansi Patel) Writing of sections of the Final Report not listed above as considered fair

References to works investigating the state-of-the-art approaches

- Abdalla, A., Alashwal, H., Crouse, J., El Halaby, M., & Moustafa, A. (2019). The Application of Unsupervised Clustering Methods to Alzheimer's Disease. *Frontiers in Computational Neuroscience*, 13, 31.
<https://dx.doi.org/10.3389%2Ffncom.2019.00031>
- Auras, H., Merouani, H. F., & Refai, A. (2016). Maintenance of a Bayesian network: application using medical diagnosis. *Evolving Systems*, 7(3), 187-196. doi: 10.1007/s12530-016-9146-8.
https://casls-primo-prod.hosted.exlibrisgroup.com/permalink/f/1ed24l6/TN_cdi_cross_ref_primary_10_1007_s12530_016_9146_8
- Govrdhan, A., Kavihta Rani, B., & Srinivas, K. (2010). Applications of Data Mining Techniques in Healthcare and Prediction of Heart Attacks. *International Journal on Computer Science and Engineering*, 2(2), 250-255.
https://www.researchgate.net/profile/Dr_Govardhan/publication/49617135_Applications_of_Data_Mining_Techniques_in_Healthcare_and_Prediction_of_Heart_Attacks/links/5435156d0cf2dc341daf6a8e.pdf