Project Proposal

Title / Team members

Private Healthcare Services Performance Data Set

Introduction (e.g. which purpose or mining issue):

We intend to explore data mining from the Kaggle deep learning repository as an approach to the market prediction for private health services. We aim to use the data to build and test a model for data mining to predict demand for private healthcare services in Ghana. This data could be useful for finding pressure points in the needs of the private health sector and therefore could be very beneficial for government and individual private health facilities.The data will aid in government as well as institutional planning to manage the ever-growing pressure in the healthcare facilities in the country.

Descriptions of Data (e.g., descriptive stats and characteristics)

As part of the Citizen Data Science initiative, this data set is given to capture and provide reasonably clean data (which is a challenge in these regions) to promote the application of data science in Ghana and other regions at the beginning of their data analytics learning process. So your assistance is welcome/This dataset includes a list of medical centers in Ghana, which helps one to obtain new knowledge of the health system of the country by researching it. There are 8 attributes, 3757 instances, and 3 target variables in the dataset, and we will. For our classification mission, mapping the continuous target variable longitude and latitude to regular letter grades ownership, the data has both continuous and discrete variables and is reasonably clean and ready for review.

Preliminary Strategy (e.g., formulation of questions, potential techniques/models)

As a supervised learning classification problem, we are formulating our problem. We plan to explore all possible Weka classifiers, but we will pay particular attention to linear regression for the role of classification and random forest for the significance of the function. We are not limited to linear approaches and may also investigate neural networks. To test and compare our models, we will use precision metrics such as accuracy, sensitivity, specificity, F1 score, etc.

Timeline

Week Five:

* Review of exploratory data
* Develop compound characteristics
* Prepare details for review

Week 6:

* For the article, write and compile data descriptions
* Run Weka classifiers for non-neural networks
* Compare different models efficiency

Week Seven

* Continue research using neural networks on Weka

Week Eight

* Start collecting report findings
* Compare outcomes to known benchmarks

Week Nine

* Complete report
* Proofread for errors
* Submit report.

**References**

Durairaj, M., & Ranjani, V. (2013). Data mining applications in healthcare sector: a study. *International journal of scientific & technology research*, *2*(10), 29-35. <https://www.academia.edu/download/45634954/Data-Mining-Applications-In-Healthcare_2.pdf>

Milovic, B., & Milovic, M. (2012). Prediction and decision making in health care using data mining. *International Journal of Public Health Science (IJPHS)*, *1*(2), 69-78. <https://www.academia.edu/download/33572666/1380-4163-1-PB_(1).pdf>

Tomar, D., & Agarwal, S. (2013). A survey on Data Mining approaches for Healthcare. *International Journal of Bio-Science and Bio-Technology*, *5*(5), 241-266. <https://www.researchgate.net/profile/Rathish_Nair2/post/Is_there_any_plain_text_dataset_available_in_medical_domain/attachment/59d628b679197b80779871d9/AS:331873965363200@1456136327356/download/A+survey+on+Data+Mining+approaches+for+Healthcare.pdf>