Analysis and Visualization of Metabolic Syndrome

Using

Self Organizing Maps (SOM)

Peter Kamita Kihato

A thesis submitted in fulfillment for the degree of Doctor of Philosophy in Electrical and Electronic Engineering in the Jomo Kenyatta University of Agriculture and Technology

2009

ABSTRACT

In recent times, a lot of research has been going on in the field of nervous systems with a view of grasping and utilizing the acquired knowledge in the area of artificial intelligence. One of the branches of science inspired by the functioning of the brain is artificial neural networks. Self Organizing Maps (SOM) falls under artificial neural networks, and can be viewed as a visualization tool that projects high-dimensional dataset onto a two-dimensional plane thereby simplifying the complexity of the monitored data. The simplification in effect discloses much of the hidden details for easy analysis, clustering and visualization, but still preserving the details of original data. The pioneer of SOM algorithms, T. Kohonen, developed plane or flat SOM data mining tool. The tool has drawbacks in that it does not consider the neighborliness or relationship between the nodes appearing at the corners and edges of the lattice. The clusters formed at these regions have no similarity. In this research improved SOM tools Torus and Spherical that overcame the flat SOM drawbacks were developed.

One of the threatening trends of human health in recent years has been metabolic syndrome. Metabolic syndrome is a cluster of conditions that occur together resulting in simultaneous health disorders related to ones metabolism. Such disorders as obesity, particularly around the waist, elevated blood pressure, elevated level of the blood fat (triglycerides) (TG), low level of high-density lipoprotein cholesterol((HDL)) and resistance to insulin (a hormone that helps to regulate the amount of sugar in the body). The disorders are taken as parameters (variables) affecting a healthy system. Having one component of metabolic syndrome means one is more likely to have others. The more components you have, the greater the risks to ones health.

The developed data mining tools were therefore, subsequently used to analyze and visualize metabolic syndrome as a risk to human health. The dataset parameters were Body Mass Index (BMI), High Blood Pressure (HBP), Blood Glucose (GLU), TG, Low Blood Pressure (LBP) and HDL. Using the developed Torus and Spherical SOM, real health data (4007 females and 2450 male test data) was used in the simulation. The contribution each risk (parameter) had on the syndrome was analyzed. Combination of parameters and their priority to induce the syndrome risk were also investigated. Also investigated using the same tools were the probable causes of the syndrome to both male and female examinees. The results obtained from the analysis compared very well with those diagnosed by the physicians thereby validating the Torus and spherical SOM. Referring to the sampling done on examinees, the ones diagnosed to be metabolic by the physicians were also found to be metabolic using the developed software. The developed simulators even reviewed trends the physicians could not have obtained at a glance. Moreover, after identifying the dominant parameters that contribute to the syndrome risk, specific (software) tools were developed to evaluate metabolic syndrome more accurately, particularly focusing on these main contributors. The developed tools were further used to formulate future trends the parameters may follow for a particular examinee. The metabolic SOM tools developed, display the metabolic syndrome risks in percentages with the most risk given 100 mark-points (MK). Examinees can equally observe the risk status they may be in. The developed tools can even predict the risk status the examinee may be in if the observed trend is not corrected through medical, psychological or physical means in good time. The risk factors were further analyzed based on age. Using spherical SOM with the simulated data moderated to read metabolic points, the age cluster trends were formulated from which it was observed that each cluster responded differently towards the syndrome risk factors.

Analysis and visualization approach to metabolic syndrome developed here has initiated a different concept of understanding and appreciating the sources of the syndrome. The visualization tools are very handy for development of the trends the risk parameters may be taking. While the actual definition of metabolic syndrome may vary for the physicians, the clustering that occurs after training SOM becomes a useful map to aid the diagnosis of the examinees. Component maps generated from the trained SOMs showing how each risk parameter affects the overall metabolic map, become very helpful to the physician since dominant risk parameters become known. Furthermore, using the resulting trend maps, physicians are able to monitor the trends the risk parameters are taking.

The developed tools become an added opinion to the physician diagnosis. The examinees are themselves advantaged by the fact that SOM tools are self-explanatory maps and therefore they can observe the risk levels and the parameters causing them to be in the position their measured parameters have mapped them. It should however be noted that physician's comment need to be considered as the professional opinion to this form of SOM application. With this in mind, the tools were deliberately developed under constant consultations with physicians. The obtained results were precise and in agreement with the interpretations from the physicians, who are the experts. From the analyses TG, HBP and BMI were found to be the highest risk factors to metabolic syndrome. Age clustering analysis isolated HBP as the most dominant risk factor. Finally, it is notable that the physician expert advice coupled with knowledge gained from examinee's interpretation of the maps, become an enhanced healing process and hence a quickened recovery period.