Using data mining technique to predict breast cancer survivability

Yu-Feng Lu, Chi-Ming Chu
School of Public Health, National Defense Medical College, Taipei, Taiwan

According to the American Cancer Society data, there were 211,240 female and 1,690 male invasive breast cancer new cases diagnosis in 2005. There were 58,490 carcinoma in situ new cases occurred per year; breast cancer is in the second place of the frequent cancer for women worldwide. According to statistic data in 2005, 1339 people died for breast cancer in Taiwan in 2003 (the death rate is 12.04 out of a hundred thousand), which is in the fourth place of women dying cancer. 5067 cases of breast cancer occurred in 2001 were diagnosis. In spite of the death rate of breast cancer is getting less year over year, it is still in the second place of women cancer.
The effective pre-screenings are considered great contributions to the death rate in developed countries getting less significantly. Breast cancer can be found early and secured through effective predictable tools and its characteristic of longer period before getting worse. However disease sufferings have close relationship with the dangerous factors, Logistic Regression is used to do predictions of the past diseases in statistics. In medical science field, we hope can use artificial intelligence exploring technology to find more possible causes in function to help to contribute to the higher correct rate of diagnoses and form a predictable health education model. It can be more positive and effective compared to the current disease secure model, and it is also helpful to the development prediction medical.

The purpose of the study is using America cancer register dataset（Surveillance Epidemiology and End Results，SEER）and Taiwan cancer register dataset to explore the use of artificial intelligence technology and the prediction model of breast cancer in America and Taiwan; moreover, to discuss the relationship between breast cancer risk factors and survival status.

Literatur

[1] Abu-Hanna, A. and P. J. Lucas (2001). “Prognostic models in medicine. AI and statistical approaches.” Methods Inf Med 40(1):1-5
[2] Burke, H.B. and D.E. Henson (1993). “The American Joint Committee on Cancer. Criteria for progostic factors and for an enhanced prognostic system.“ Cancer 72(10): 3131-5.