Abstract

The paper builds a prediction model of 30-day mortality risk in patients with CHF. Least absolute shrinkage and selection operator (LASSO) was applied to select the significant features. Three data mining techniques, namely, decision tree, logistic regression, and AdaBoost.M1 algorithms were used to predict mortality risk in patients with CHF. A case study was conducted using data (January 2012 to December 2014) from a community hospital in Upstate New York and a comparison among the three predictive algorithms was performed. The primary measure for comparing the performance of the prediction algorithms were area under the receiver operating characteristic curve (AUC), specificity, sensitivity, and accuracy. The results show that logistic regression model resulted in an accuracy of 84.85% with the specificity, sensitivity and AUC of 1, 0.44 and 0.72, respectively. On the other hand, the decision tree algorithm resulted in an accuracy of 75.76% with specificity, sensitivity and AUC of 0.83, 0.56 and 0.69, respectively, whereas the AdaBoost.M1 algorithm resulted in 81.82% accuracy with specificity, sensitivity and AUC of 0.56, 0.54 and 0.55, respectively. This research concludes that the logistic regression model fitted with LASSO outperforms AdaBoost.M1 and decision tree in terms of both AUC score as well as predictive accuracy. Accurate prediction of 30-day mortality based on this research can be useful in risk stratification, individualized treatment, and patient management.

Citation

Khurjekar, N., Aladeemy, M., Chou, C., Shan, X., Poranki, S. Khasawneh, M. T., and Srihari, K.,

"30-day mortality prediction of patients with congestive heart failure (CHF), Proceedings

of the Industrial and Systems Engineering Research Conference (ISERC), Anaheim, CA,

May 21-24, 2016.