Use of the Kernel Support Vector Machine for Prediction of Need for Admission and Time for Disposition among Simulated Emergency Department Patients During Disaster Exercises

Franc JM†‡, Verde M‡.

†The University of Alberta, Edmonton, Alberta, Canada
‡L'Universita' degli Studi del Piemonte Orientale, Novara, Italy

INTRODUCTION: In times of increased patient numbers, such as mass casualty incidents or other disasters, hospital emergency departments often need to increase their surge capacity. Predicting the need for surge capacity is difficult, and more accurate predictions of which patients will need admission to hospital and how long they will spend in the department may be helpful. Traditionally, predictions of patient flow are often based on the patient's triage score; this method of prediction is often rudimentary. Computerized machine learning may offer a more accurate prediction rule. Using a large dataset of simulated patients, the present study investigates the use of the Kernel Support Vector Machine as a more accurate predictor of patient flow.

METHODS: The present study investigates methods to predict two response variables: (1) need for admission and (2) time for disposition for simulated disaster patients. Two decision tools are compared for predicting the need for admission: a simple classifier based on triage score and a Kernel Support Vector Classifier. For prediction of time to disposition, again two tools are compared: a simple model based on triage score, and a Kernel Support Vector Regression. The null hypothesis that simple decision tools are equal to Kernel based tools was tested against the alternative hypothesis that the methods are not equal. The data was obtained from 65 individual simulations of the SurgeSim emergency department surge capacity simulator (SurgeSim, Edmonton, Alberta, Canada). The original SurgeSim dataset contains 6887 observations (rows) and 266 predictors (columns). Unfortunately, the dataset is complicated. It contains of observations of integers, numeric, character, text, and image references and has a complex dependency structure. As such, we limited the study to 66 columns that are most likely to contain information that is both available during a disaster situation and likely to influence the response variables. The dataset was divided into a training and test set by randomly selecting 1000 observations for the test set and leaving the remaining in the training set. Modeling used the `kvsm` function from the `kernlab` package of the R statistical computing language (R foundation for statistical computing, Vienna, Austria).

RESULTS: Using a simple triage code based prediction rule for admission gave an overall accuracy of 0.325 and a recall of 0.25. Use of the simple triage code based predictor for time to disposition was unsuccessful, as Pearson's product-moment correlation fails to show significant correlation between the observed and predicted values of time to disposition(Corelation Coefficient=0.169, 95% CI -0.05 to 0.37, p=0.13). Conversely, overall accuracy for the kernel support vector machine predictor of need for admission was 0.725, and recall was 0.6. Likewise, the kernel support vector machine predictor of time to disposition showed significant correlation by Pearson's product-moment correlation (Corelation coefficient 0.25, 95% confidence interval 0.03 to 0.044, p=0.02).

CONCLUSIONS: Prediction rules for need for admission and time to disposition decision based on the kernel support vector machine appear to be superior to simple decision rules based on triage code. These prediction rules may be valuable in timea when emergency departments require increased surge capacity. Although the present model is complex, it is based on freely available statistics software and could potentially be implemented in a computerizes patient tracking system.