AANS/CNS Joint Cerebrovascular Annual Meeting
January 22–23, 2018
Los Angeles, CA

Evaluating the Utility of Machine Learning Classifiers in Predicting Aneurysm Rupture Status

Kiran G Thomas BS; Abhinav Changa BS; Gerald Riccardello, Jr. ME; Marco S Amaral BS; Nitesh Damodara; Max Roman PhD; I. Paul Singh MD; Chirag D. Gandhi MD, FACS; Charles J. Prestigiacomo MD, FACS

UC Health

Introduction

Previous studies have demonstrated the usefulness of binary logistic regression equations in classifying the rupture status of cerebral aneurysms. However, the prediction accuracies that have been obtained remain short of the threshold necessary to become clinically relevant. Conventional approaches to improve accuracy include increasing the sample size, model complexity and variable considerations. This study aims to explore the utility of an alternative method. We employ a simple machine learning classifier on sample aneurysm data to evaluate the effectiveness of such a technique in predicting aneurysm rupture status.

Methods

Based on sample data from the literature, we created two datasets. Each dataset consisted of generated time-averaged wall shear stress (TAWSS) values and aspect ratios (AR) with corresponding ruptured or unruptured statuses. A support vector machine was trained and then tested using each dataset respectively. All data analyses were performed in MATLAB.

Results

The means and standard deviations of the generated data reflected values obtained from the literature. The training phase identified 4 support vectors for the unruptured population and 5 support vectors for the unruptured population. These support vectors were then used to identify a decision boundary which accurately separated the aneurysm populations by rupture status. During the testing phase, this support vector machine was able to classify 95% of the test data correctly.

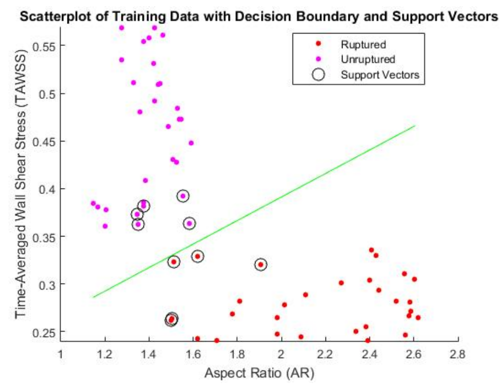
Conclusions

Our results underscore the potential of machine learning classifiers. Further study should involve retrospective, patient-specific data and should additionally aim to include a larger sample size and additional variable considerations.

Learning Objectives

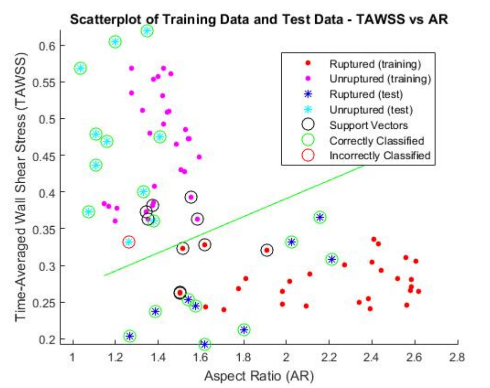
By the conclusion of this session, participants should be able to: (1) understand the theoretical underpinnings of machine learning algorithms; (2) evaluate the effectiveness of such algorithms when applied to test data; (3) consider potential applications for such algorithms in subsequent data analyses.

Scatterplot of Training Data with Decision Boundary and Support Vectors



In this figure, the results of the training phase can be observed. The relevant support vectors are illustrated and the decision boundary based on these support vectors is also delineated.

Scatterplot of Training Data and Test Data - TAWSS vs AR



In this figure, the results of the testing phase can be appreciated. The test data is plotted along with the training data, support vectors and decision boundary. The accuracy of the algorithm's classification is also demarcated.