1114-92-221   **Phuong Hoang*** (`phoang@ncsu.edu`), Department of Mathematics, Campus Box 8205, North Carolina State University, Raleigh, NC 27695, and **Hien T Tran** (`tran@ncsu.edu`), Department of Mathematics, Campus Box 8205, North Carolina State University, Raleigh, NC 27695. *Supervised learning for medical diagnostic decision support system: Hepatitis C and breast cancer diagnosis application.*

Many decision support applications in healthcare depend heavily on early detection and accurate diagnosis. Physician diagnostic performance can be improved with a reliable decision support system. In this talk, we will discuss the machine learning approach to improve the accuracy of these diagnostic systems. Machine learning techniques from classification to model evaluation will be introduced including Support Vector Machine (SVM), Receiver Operating Characteristic (ROC) curve and k-fold Cross Validation. In our initial effort, we test our approach using hepatitis C and breast cancer datasets from the UCI Machine Learning database. It is noted that most datasets in disease diagnostic are imbalanced, i.e., one class heavily outnumbers the others (for example, detecting credit card fraud or email spam). By using a modified SVM with weighted cost, we were able to obtain over 80% accuracy in both sensitivity and specificity.

Similar machine learning approach was previously employed in domain beyond health care. In particular, in collaboration with the MIT Lincoln Laboratory, we built a pitch type prediction model using a database created by Sportvision's PITCHf/x pitch tracking system and obtained 78% overall prediction accuracy. (Received August 28, 2015)