



HUMANITY  
CENTERED  
ROBOTICS  
INITIATIVE



## Rich Caruana Microsoft Research

### TECH TALK: "Do Deep Nets Really Need To Be Deep? Do Convolutional Nets Need To Be Convolutional?"



**Wednesday, October 26, 2016**  
**3-4pm**  
**CIT 506**

*Abstract:* Deep neural networks are the state of the art on problems such as speech recognition and computer vision. Using a method called model compression, we are able to train shallow nets to learn the complex functions previously learned by deep nets and achieve accuracies previously only achievable with deep models while using the same number of parameters as the original deep models. On the TIMIT phoneme recognition and CIFAR-10 image recognition tasks, shallow nets can be trained that perform similarly to complex, well-engineered, deeper convolutional architectures. The same model compression trick that we're using to examine if depth and convolution are important is also used to compress impractically large deep models and ensembles of large deep models down to small- or medium-size deep models that run more efficiently on mobile devices or servers.

**Rich Caruana** is a Senior Researcher at Microsoft Research. Before joining Microsoft, Rich was on the faculty in the Computer Science Department at Cornell University, at UCLA's Medical School, and at CMU's Center for Learning and Discovery. Rich's Ph.D. is from Carnegie Mellon University, where he worked with Tom Mitchell and Herb Simon. His thesis on Multi-Task Learning helped create interest in a new subfield of machine learning called Transfer Learning. Rich received an NSF CAREER Award in 2004 (for Meta Clustering), best paper awards in 2005 (with Alex Niculescu-Mizil), 2007 (with Daria Sorokina), and 2014 (with Todd Kulesza, Saleema Amershi, Danyel Fisher, and Denis Charles), co-chaired KDD in 2007 (with Xindong Wu), and serves as area chair for NIPS, ICML, and KDD. His current research focus is on learning for medical decision making, deep learning, and computational ecology.

Host: Michael Littman/HCRI