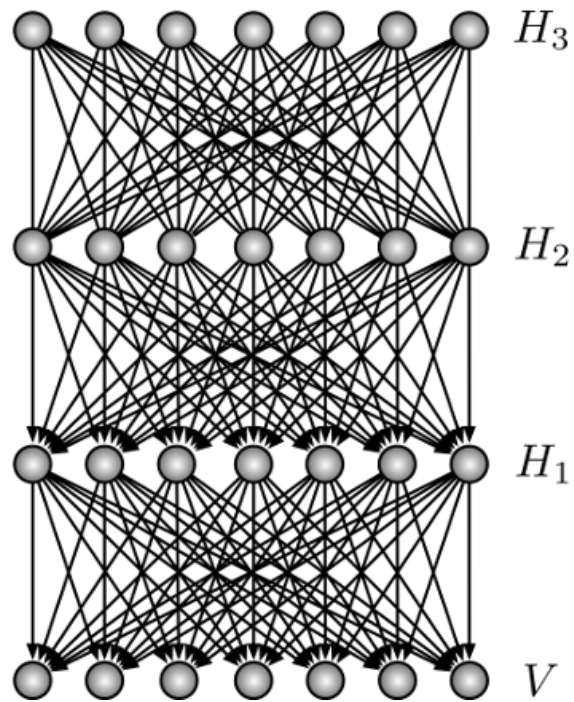# Computer Science 577 Notes
Machine Learning

Mendel C. Mayr

January 23, 2015

Contents

1 Decision Tree Learning

Information theory: entropy and information gain

(i) Entropy $= H(X) = -\sum_{x \in X} P(X = x) \log_2 P(X = x)$

(ii) Conditional entropy $= H(Y|X) = -\sum_{x \in X} P(Y|X = x)$, where
$H(Y|X = x) = \sum_{y \ inY} P(Y = y|X = x) \log_2 P(Y = y|X = x)$

(iii) Mutual information (information gain) $= I(X, Y) = H(Y) - H(Y|X)$