A decision tree can be built to perfectly represent f(X). A general example of a ‘perfect’ tree would have a depth of ‘d’ – each level of the tree would involve splitting based on one attribute. Since our training data includes all 2^d values of f(X), and f(X) is not noisy – it returns the same value every time – our tree could, in theory, split in an arbitrary order and end up with 2^d leaf nodes which would correctly predict f(X) for any of the 2^d values of X.