

Example problem: "Given this data, a friend has a house 750 square feet - how much can they be expected to get?" What approaches can we use to solve this? Straight line through data Maybe \$150 000 Second order polynomial Maybe \$200 000 One thing we discuss later - how to choose straight or curved line? Each of these approaches represent a way of doing supervised learning What does this mean? We gave the algorithm a data set where a "right answer" was provided So we know actual prices for houses The idea is we can learn what makes the price a certain value from the training data The algorithm should then produce more right answers based on new training data where we don't know the price already i.e. predict the price We also call this a regression problem Predict continuous valued output (price) No real discrete delineation Another example Can we define breast cancer as malignant or benign based on tumour size