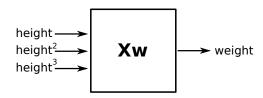
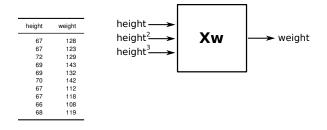
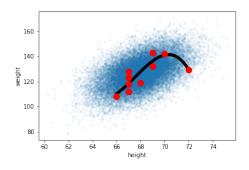
gorithm.

35 Example (Feature Engineering Height vs Weight)
In order to improved the prediction accuracy of weight using height, we can include two new features: height² and height³.



36 Example (Height vs Weight Feature Engineering)





37 Example (Hyper-Parameters)

Consider the problem of predicting a child's weight, y, from the child's height, x. Assume we have created a series of engineered feature, x^k , $k=2,3,\ldots,d$ so that

$$\hat{y} = b + w_1 x + w_2 x^2 + \dots + w_d x^d.$$

The parameter d is called a **hyper-parameter**. The parameters b, w_1, w_2, \ldots, w_d are selected to minimize training MSE by solving the normal equations. The hyper-parameter d, on the other hand, is selected to minimize the test MSE. In the Lesson (Feature Engineering), the optimal value for the hyper-parameter d is d=1.

38 Example (Stock Prices)

Can past stock prices be used to predict future stock prices?

The file AdjustedClosingPrices.csv⁵ contains data on 467 U.S. stocks over the course of 4173 days. Train a linear regression neural network that uses yesterday's and today's stock prices to predict tomorrow's stock prices. Train the network using stock price data for IBM. Normalize the data by computing log price ratios as shown below:

$$\mbox{log price ratio } = \ln \left(\frac{\mbox{price today}}{\mbox{price yesterday}} \right).$$

Evaluate the accuracy of your trained neural network.

39 Example (Iris Dataset)

R. A. Fisher's iris data set is one of the oldest, most used examples in pattern recognition. The data set contains three class (50 instances each) of the iris plant. Each class corresponds to one of three plant species: Setosa, Versicolour and Virginica. Plant attributes consist of four measurements (in centimeters) of plant leaves: sepal length and width and petal length and width.

⁵Courtesy Dylan Vener. The data was downloaded from Yahoo Finance.