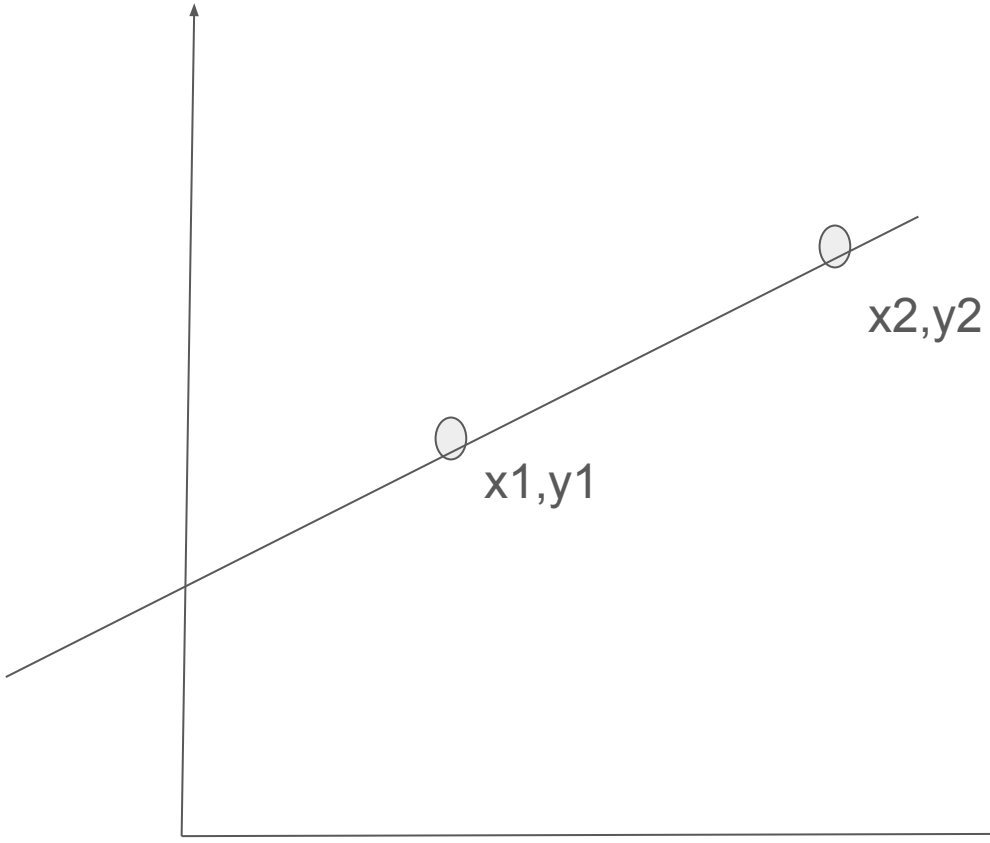


1. Calculate slope of the line m = (y2-y1)/(x2-x1)
2. y=**m**x+**c**
3. Calculate the value of c

Y = 5X+6 (for a given set of data points)

Y = 3X-6 (for another set of points)

y = f(X1, X2, X3, X4….)

X1 = Weight
X2 = height
X3 = Blood Sugar Levels…

y = mX + c -> Simple Linear Regression

y = m1*X1 + m2*X2 +... + c - Multiple Linear Regression

Y = 5x1 + 6x2 + 7

Y = f(x1, x2)

x2 = f(x1) Not independent

Objective of LR = Min (Actual - predicted)^2

Predicted value = m1x1+m2x2+m3x3… + c

Fit Transform - > Two steps

1) Fit -> Learn the mean and sd for every column
2) Transform -> Apply the z transform (actual - mean)/sd

Transform - > 1 step

Training Set -> Apply fit_transform, Test Set -> Transform

We only learn anything from the training set. We do not use the test for anything.

Overfitting (error on Test set is quite high)
RMSE on Training set is very low < 1, RMSE on Test set is high > 5

Underfitting (model has not learnt anything from the training data)
RMSE on Training set is high > 5, RMSE on Test set is also high > 5

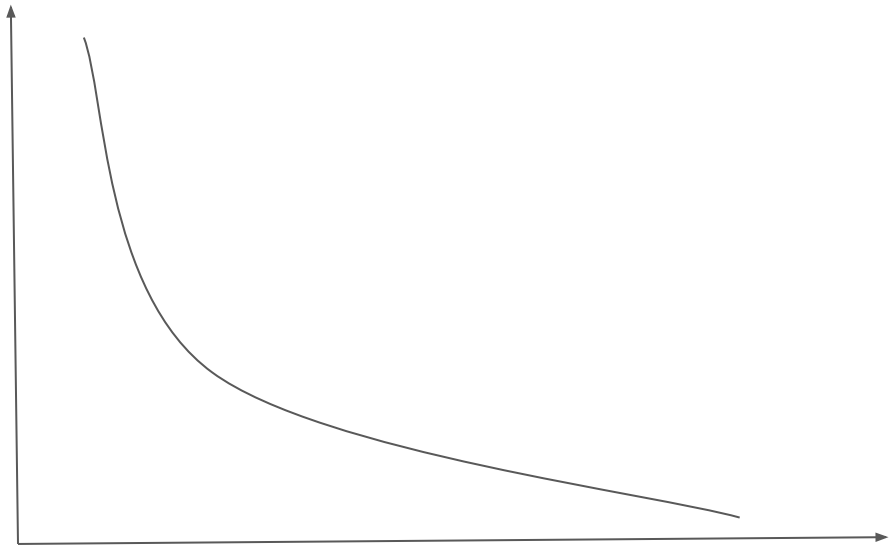
Training RMSE -> 1.6, Test RMSE -> 2.3

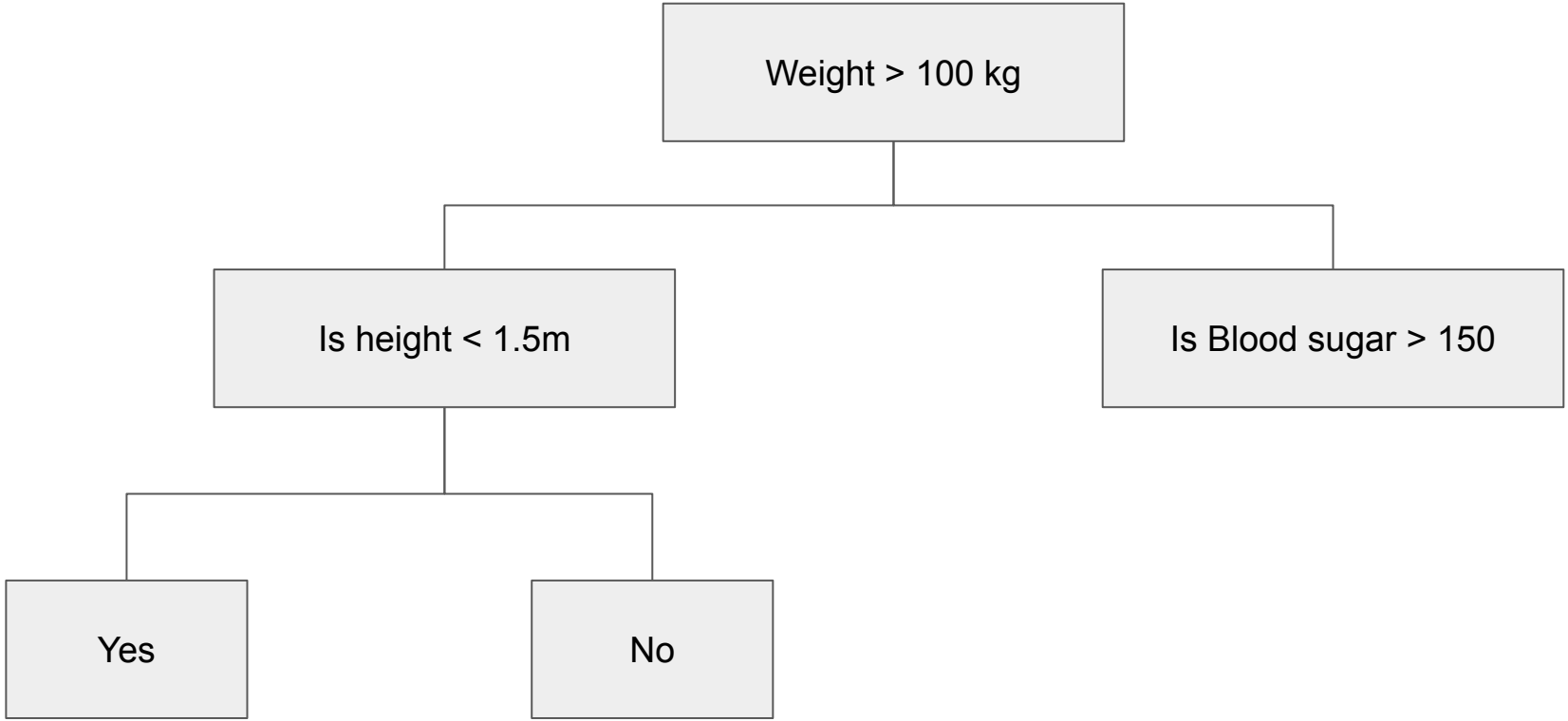
Regularization
Objective function of Linear Reg model =

Min (Actual - predicted)^2 + penalty term

Polynomial Regression

$y = m2.(x1)^2 + m2.(x2)^2$




Weight > 100 kg
Is height < 1.5m
Is Blood sugar > 150
Yes
No