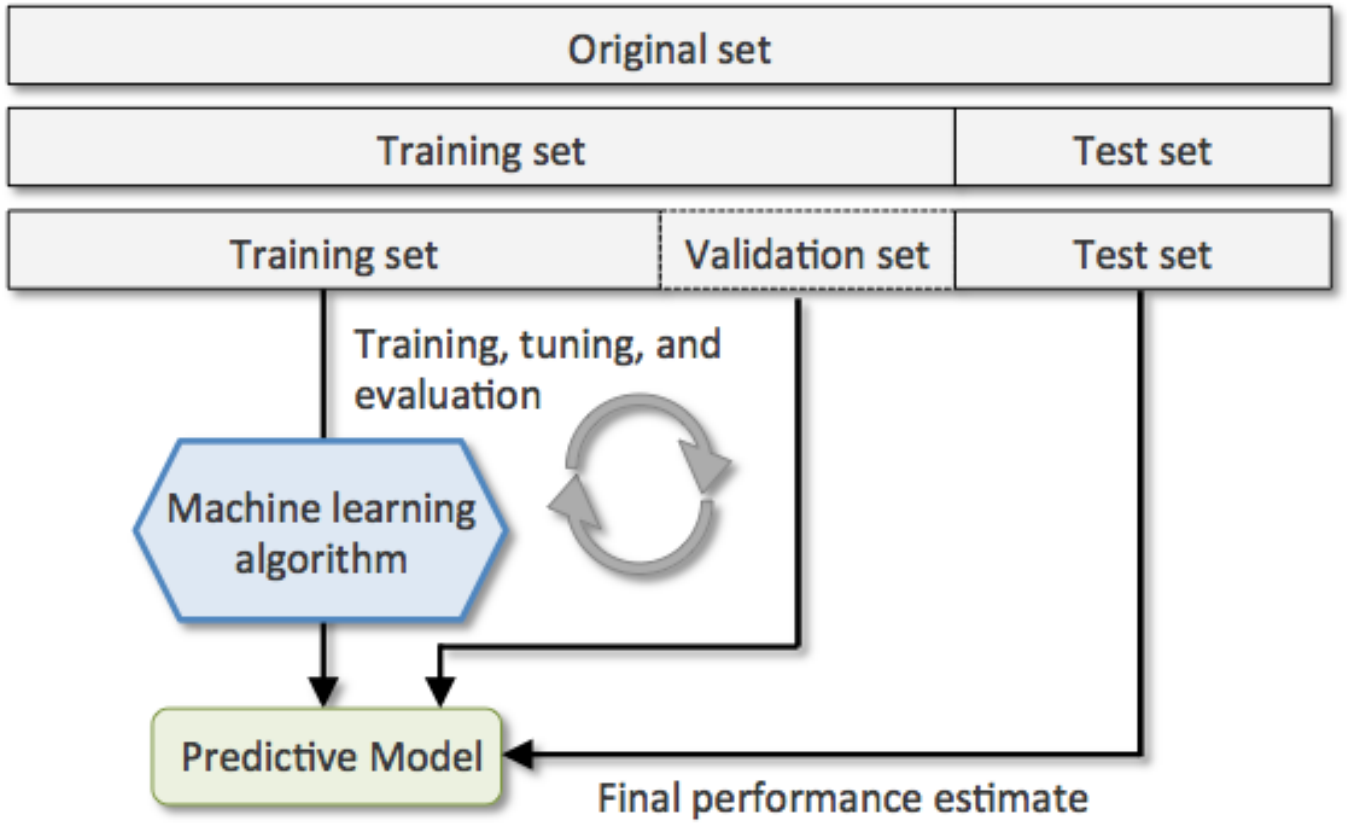
Test Train split

Test Train Split



Random Seed

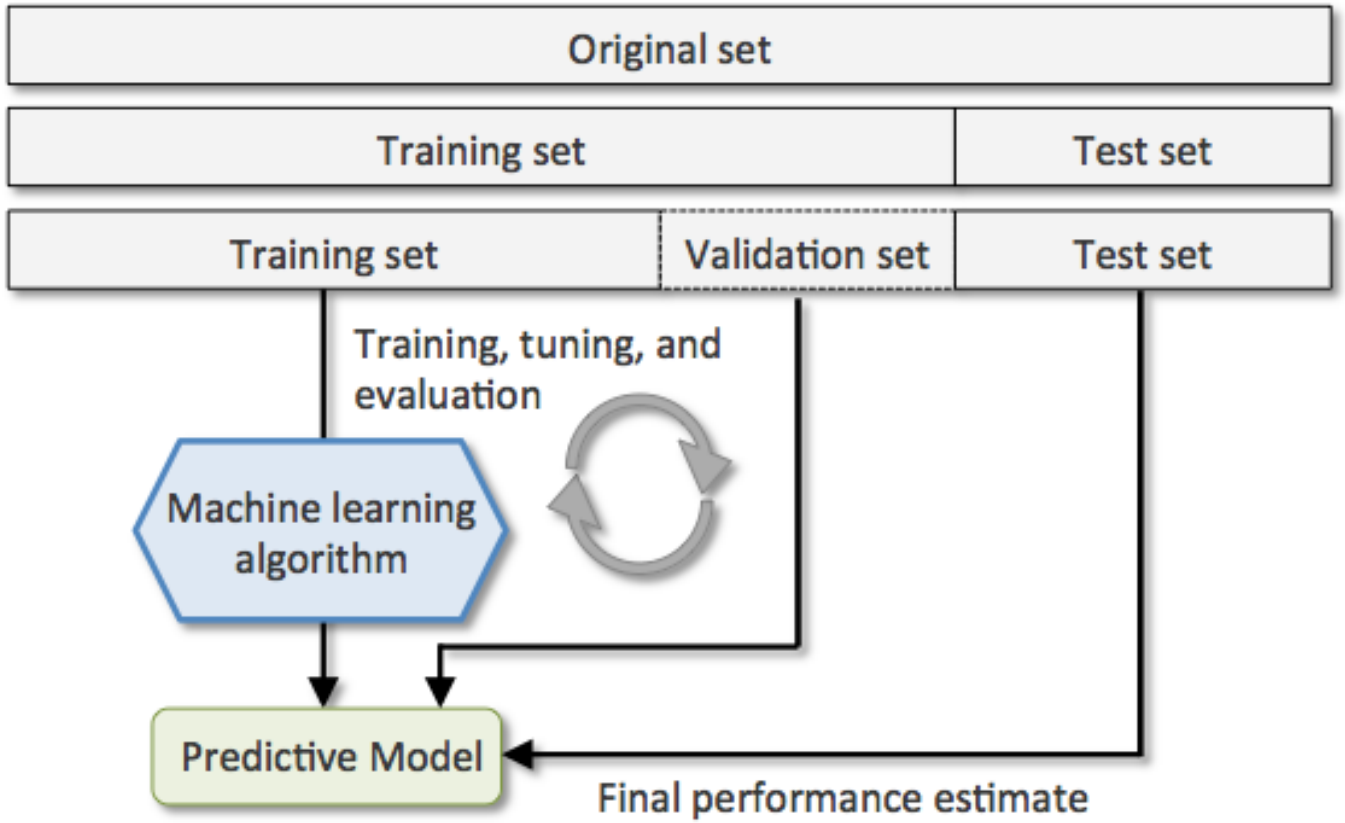
WHY

Generate same set of Random numbers while

1. Test Train Split
2. Model Training

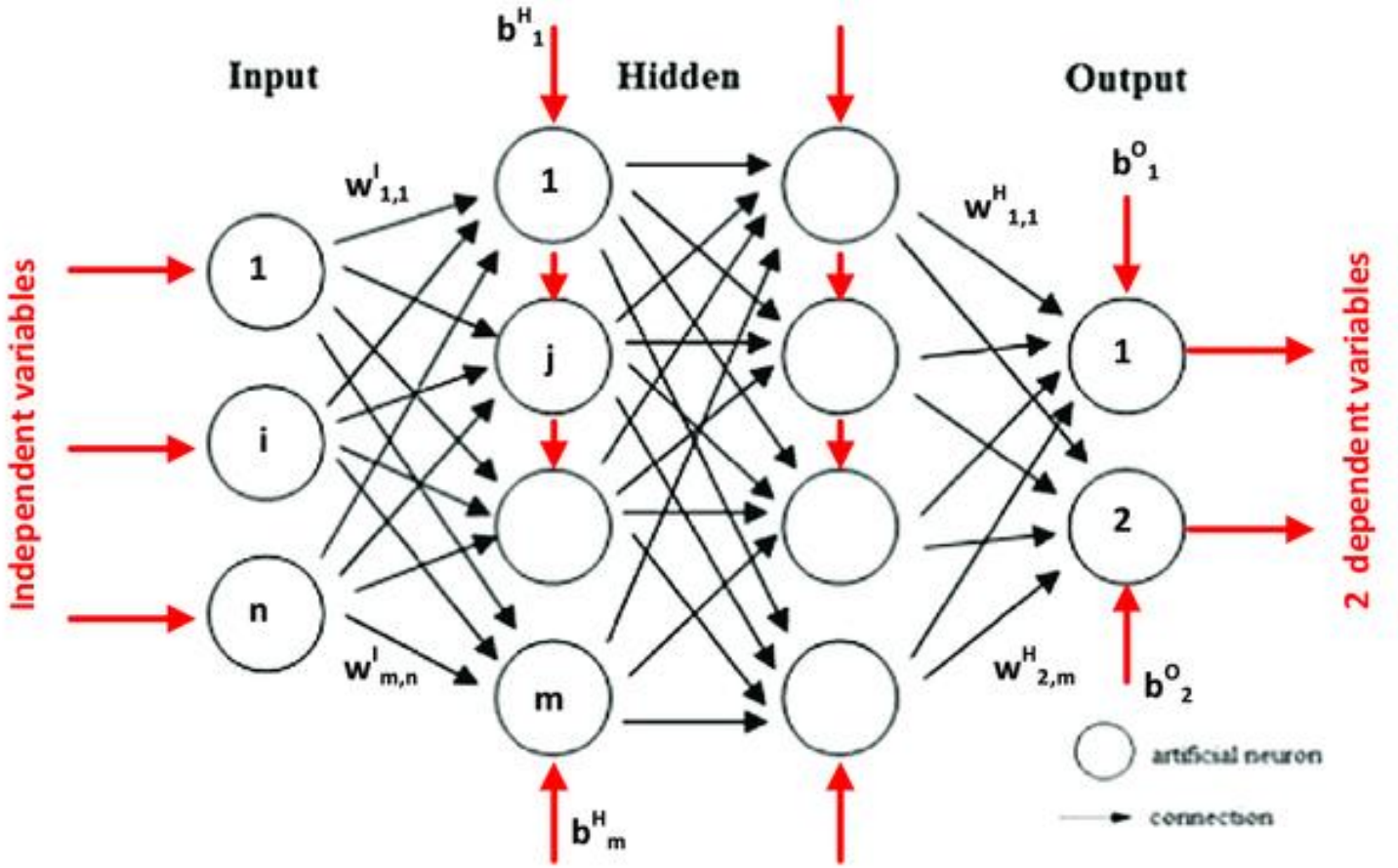
Random Seed

Test Train Split



Random Seed

Model Training



Feature Scaling

Types

Feature scaling is a method used to normalize the range of independent variables or features of data.

1. Normalization
2. Standardization

Feature Scaling

Example

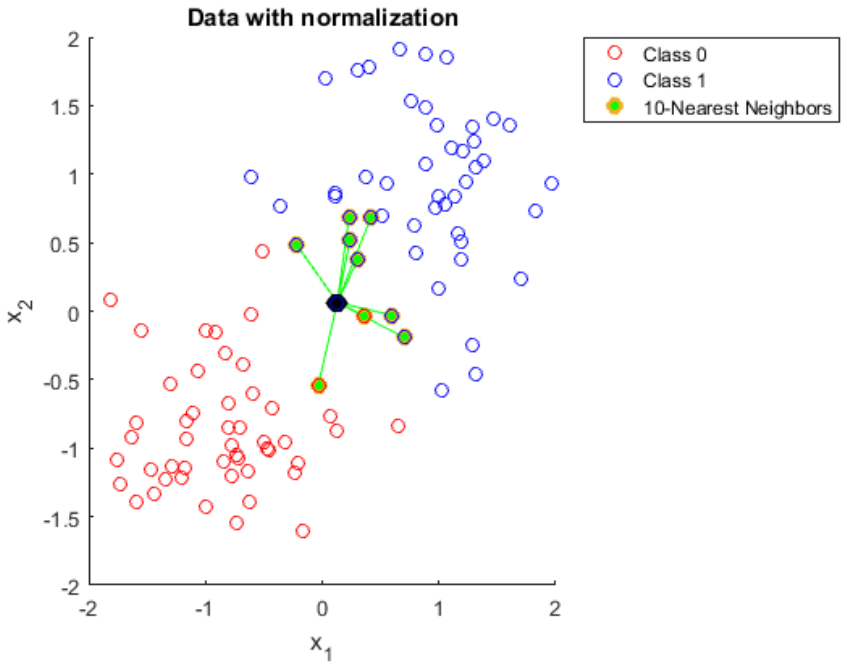
Predicting whether or not you want to give loan depending on.

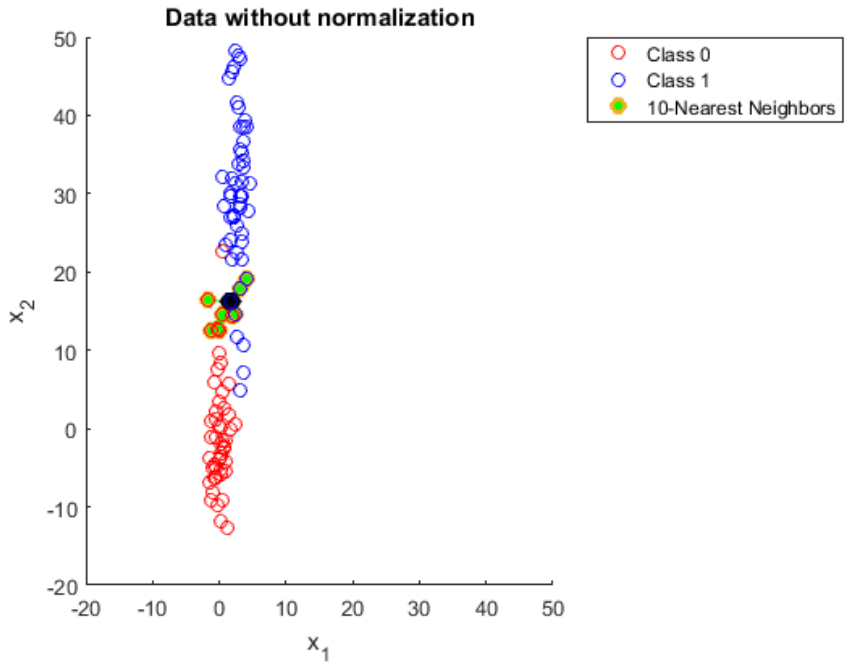
1. Age – Range 0-100
2. Income – 10,000 to 10,000,000

Feature Scaling

Age – Range 0-100

Income – 10,000 to 10,000,000



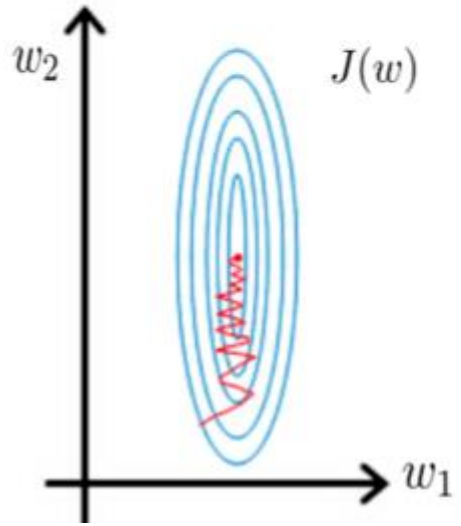


Feature Scaling

WHY

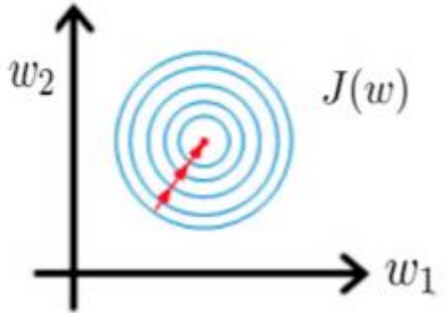
Gradient descent without scaling

$x_1 >> x_2$



Gradient descent after scaling variables

$0 \leq x_1 \leq 1$
$0 \leq x_2 \leq 1$



Feature Scaling

Normalization

$$x_{\text{norm}} = \frac{x - \min(x)}{\max(x) - \min(x)}$$

Values 2,4,6

2 will become = (2-2/(6-2) = 0

4 will become = (4-2/(6-2) = 0.5

6 will become = (6-2/(6-2) = 1

Feature Scaling

Standardization

$$x_{\text{stand}} = \frac{x - \text{mean}(x)}{\text{standard deviation } (x)}$$

Values 2,4,6 => mean =4 & SD is 2

2 will become = (2-4)/(2) = -1

4 will become = (4-4/(2) = 0

6 will become = (6-4/(2) = 1

Feature Scaling

Example

Age =20,40,60 Income = 40,000 , 80,000 & 120,000

Age 20 will become = 0 in norm. and -1 stand.

Age 20 will become = 0.5 in norm. and 0 stand.

Age 60 will become = 1 in norm. and 1 stand.

Income 40,000 will become = 0 in norm. and -1 stand.

Income 80,000 will become = 0.5 in norm. and 0 stand.

Income 120,000 will become = 1 in norm. and 1 stand.