

An abstract graphic on the left side of the slide. It features a network of white dots connected by thin white lines, forming a complex web. Overlaid on this network is a large blue gear icon with a white outline. To the right of the gear is a blue megaphone icon, also with a white outline. The background is a solid dark blue. There are some yellow lines at the top left and some small blue triangles scattered around the network.

# Data Preprocessing



# The Machine Learning Process

# The Machine Learning Process



## Data Pre-Processing

- Import the data
- Clean the data
- Split into training & test sets
- Feature Scaling



## Modelling

- Build the model
- Train the model
- Make predictions



## Evaluation

- Calculate performance metrics
- Make a verdict



# Training Set & Test Set

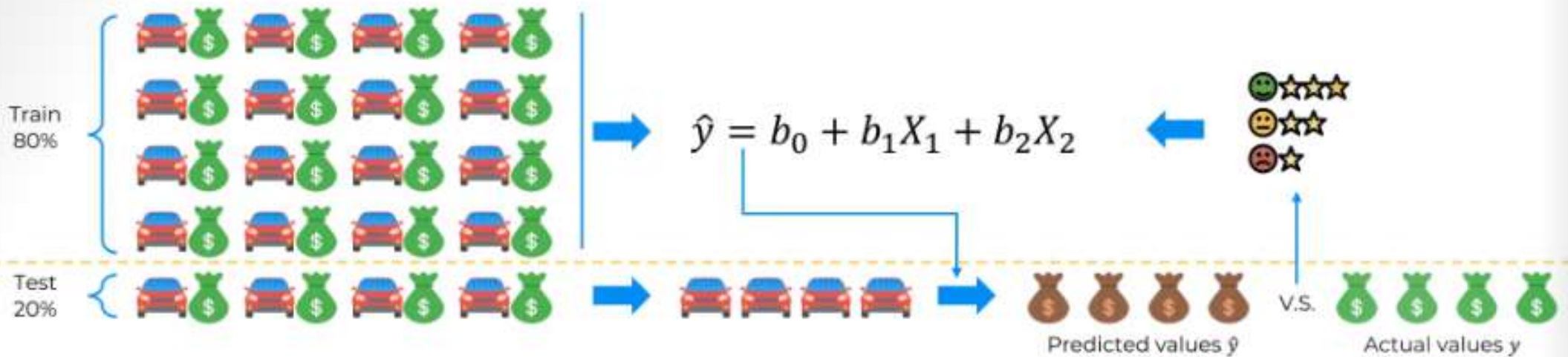




# Training Set & Test Set



~



An abstract graphic on the left side of the slide. It features a central blue gear icon with a white outline, surrounded by a network of blue triangles of various sizes, some solid and some outlined. The triangles are connected by thin black lines, creating a mesh-like structure. The overall color palette is blue and white, with some yellow lines at the top.

# Feature Scaling

# Feature Scaling



X1	X2	X3	X4
\$ 179.43	56.784	34.6181	3.55
\$ 641.87	62.054	47.7306	1.692
\$ 556.30	64.13	55.596	1.559
\$ 578.47	63.377	52.7121	1.679
\$ 591.16	61.553	46.1315	1.984
\$ 242.03	58.29	39.2952	2.942
\$ 364.66	59.93	42.4628	2.494
\$ 190.68	57.271	36.2725	3.419
\$ 547.23	63.763	54.1971	1.634
\$ 359.69	59.375	41.5105	2.128
\$ 438.08	60.484	43.493	2.47
\$ 637.17	62.525	49.428	1.725



# Feature Scaling



Normalization

$$X' = \frac{X - X_{min}}{X_{max} - X_{min}}$$

[0 ; 1]

Standardization

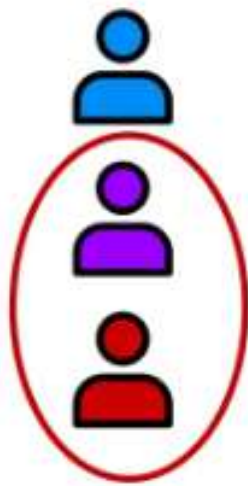
$$X' = \frac{X - \mu}{\sigma}$$

[-3 ; +3]





# Feature Scaling



70,000 \$  
10,000  
60,000 \$  
8,000  
52,000 \$

45 yrs  
1  
44 yrs  
4  
40 yrs



# Feature Scaling



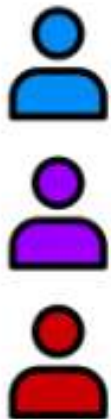
Normalization

$$X' = \frac{X - X_{min}}{X_{max} - X_{min}}$$

[0 ; 1]



# Feature Scaling



70,000 \$

60,000 \$

52,000 \$



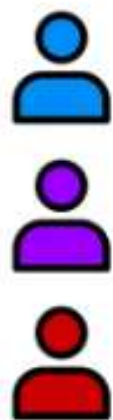
45 yrs

44 yrs

40 yrs



# Feature Scaling



1  
0.444  
0

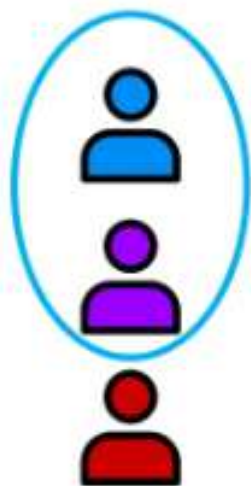


45 yrs  
44 yrs  
40 yrs





# Feature Scaling



1  
0.444  
0



1  
0.75  
0

