Task 01 Feature Transformation

(x) Price = 110, 105, 115, 120, 110, 130, 150, 100, 105

Normalization

$$x_{new} = \frac{x_i - \min(x)}{\max(x) - \min(x)}$$

$\min(x) = 100$
$\max(x) = 150$

Prices ⇒

$x_1 = \frac{110 - 100}{150 - 100} = \frac{10}{50} = 0.2$

$x_2 = \frac{105 - 100}{50} = \frac{5}{50} = 0.1$

$x_3 = \frac{115 - 100}{50} = \frac{15}{50} = 0.3$

$x_4 = \frac{120 - 100}{50} = \frac{20}{50} = 0.4$

$x_5 = \frac{110 - 100}{50} = \frac{10}{50} = 0.2$

$x_6 = \frac{130 - 100}{50} = \frac{30}{50} = 0.6$

$x_7 = \frac{150 - 100}{50} = \frac{50}{50} = 1.0$

$x_8 = \frac{100 - 100}{50} = \frac{0}{50} = 0.0$

$x_9 = \frac{105 - 100}{50} = \frac{5}{50} = 0.1$

Prices = 0.2, 0.1, 0.3, 0.4, 0.2, 0.6, 1, 0, 0.1

Standardization

$$x_{new} = \frac{x_i - x_{mean}}{\text{Standard Deviation}}$$

Standard deviation, $\sigma = \sqrt{\frac{\Sigma(x_i - \mu)^2}{N}}$

$$x_{mean} = (110+105+115+120+110+130+150+100+105)/9$$

$$= 116.11$$

$\sigma = 15.57$

$$x_1 = \frac{110 - 116.11}{15.57} = \frac{-6.11}{15.57} = -0.39$$

$$x_2 = \frac{105 - 116.11}{15.57} = \frac{-11.11}{15.57} = -0.71$$

$$x_3 = \frac{115 - 116.11}{15.57} = -0.07$$

$$x_4 = \frac{120 - 116.11}{15.57} = 0.24$$

$$x_5 = \frac{110 - 116.11}{15.57} = -0.39$$

$$x_6 = \frac{130 - 116.11}{15.57} = 0.89$$

$$x_7 = \frac{150 - 116.11}{15.57} = 2.18$$

$$x_8 = \frac{100 - 116.11}{15.57} = -1.03$$

$$x_9 = \frac{105 - 116.11}{15.57} = -0.71$$

Log transformation

$$x_{new} = \log_{10}(x_i)$$

$x_1 = \log_{10} 110 = 2.04$

$x_2 = \log_{10}(105) = 2.02$

$x_3 = \log_{10}(115) = 2.06$

$x_4 = \log_{10}(120) = 2.08$

$x_5 = \log_{10}(110) = 2.04$

$x_6 = \log_{10}(130) = 2.11$

$x_7 = \log_{10}(150) = 2.18$

$x_8 = \log_{10}(100) = 2.00$

$x_9 = \log_{10}(105) = 2.02$

Robust Scaler

$$x_{scale} = \frac{x_i - x_{median}}{x_{75} - x_{25}}$$

$IQR = x_{75} - x_{25}$

$= 120 - 105$

$= 15$

$x_{median} = 110$

$x_1 = \frac{110 - 110}{15} = 0.00$

$x_2 = \frac{105 - 110}{15} = -0.33$

$x_3 = \frac{115 - 110}{15} = 0.33$

$x_4 = \frac{120 - 110}{15} = 0.66$

$x_5 = \frac{110 - 110}{15} = 0.00$

$x_6 = \frac{130 - 110}{15} = 1.33$

$x_7 = \frac{150 - 110}{15} = 2.66$

$x_8 = \frac{100 - 110}{15} = -0.66$

$x_9 = \frac{105 - 110}{15} = -0.33$

Max Absolute Scaler

$$x_{scaled} = \frac{x}{max(x)}$$

$max(x) = 150$

$x_1 = \frac{110}{150} = 0.73$

$x_2 = \frac{105}{150} = 0.70$

$x_3 = \frac{115}{150} = 0.77$

$x_4 = \frac{120}{150} = 0.80$

$x_5 = \frac{110}{150} = 0.73$

$x_6 = \frac{130}{150} = 0.87$

$x_7 = \frac{150}{150} = 1.00$

$x_8 = \frac{100}{150} = 0.67$

$x_9 = \frac{105}{150} = 0.70$