Feature Normalisation

COMP337/COMP527 - Data Mining and Visualisation

Procheta Sen

UNIVERSITY OF LIVERPOOL

Feature normalisation: [0,1]-scaling

There are various ways to normalise (scale) a numerical features into a common scale

- **Method 1**: $[0,1]$-scaling

$$\hat{x} = \frac{x - \min(x)}{\max(x) - \min(x)},$$

  where the minimum and maximum values of the feature are computed over all training data points.

  The values of the scaled feature will now be in the interval $[0,1]$

Feature normalisation: Gaussian Normalisation

- **Method 2**: Gaussian Normalisation

$$\hat{x} = \frac{x - \mu}{\sigma},$$

where the mean (μ) and the standard deviation (σ) for the feature are computed over all training data points.

After this transformation each feature will have a zero mean and a unit variance.

Therefore, it is “easier” to compare two features, ignoring their absolute scales.