Internet of Things Assignment Advanced Feature

Description

The simple linear regression of two vectors model's the relationship between the dependant and the independent vector by estimating the model parameters from available data. In geometric terms, the parameters of the model define a linear function or straight line that maps the relationship of the independent vector to the dependant vector. Let X be the independent variable and Y be the dependent variable. The relationship is given as follows:

$$Y = mX + c$$

Here, m is the slope, i.e., the steepness of the regression line. c is the intercept, the predicted value of Y when X is 0. To find the best value for the parameters, we use the least-squares approach, i.e., find a line that minimizes the sum of the squared vertical distance between the point of the vectors and the regression line. The slope is calculated with the equation:

$$m = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^{n}(x_i - \bar{x})^2}$$

The intercept is obtained by the equation:

$$c = \bar{y} - m\bar{x}$$

Here $\bar{x}$ and $\bar{y}$ represent the mean of values in the X and Y vectors respectively.

In addition to finding the relationship of two vectors, the extent of this relationship can be found by their correlation. Specifically, the Pearson correlation coefficient is a measure of linear correlation between two vectors. This coefficient is given by the ratio of the covariance of two vectors and the product of their standard deviations. The coefficient has a value between +1 and -1, where 1 is a perfect positive linear correlation, 0 is no linear correlation, and −1 is a perfect negative linear correlation. The coefficient r_{xy} is calculated by the equation:

$$r_{xy} = \frac{cov(X,Y)}{\sigma_x \sigma_y}$$

Where σ_x and σ_y represent the standard deviation for X and Y respectively. The covariance is obtained by:

$$cov(X,Y) = Mean[XY] - Mean[X]Mean[Y]$$

Implementation

The light buffer is taken as the independent vector and the temperature buffer as the dependant vector. To find slope, the numerator of the equation is evaluated by taking the sum of the product of deviation of each value of light and temperature from their respective mean. The denominator is obtained by the sum of the square of the deviation of each value of light from its mean. The ratio of calculated numerator and denominator values gives the slope. The intercept is calculated by the difference of the mean of temperature values and the product of slope and mean of light values. The Pearson correlation coefficient is obtained by first forming an array of the product of light and temperature values and then finding the mean of this array. The covariance is then calculated and then coefficient by the ratio of covariance and product of standard deviation of light and temperature values. The output of advanced features can be observed in figure 1 and figure 2. The following formula wasn't used to calculate the coefficient as it is numerically unstable and results in the force restart of the simulation.

$$r_{xy} = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \bar{x})^2}\sqrt{\sum_{i=1}^{n}(y_i - \bar{y})^2}}$$

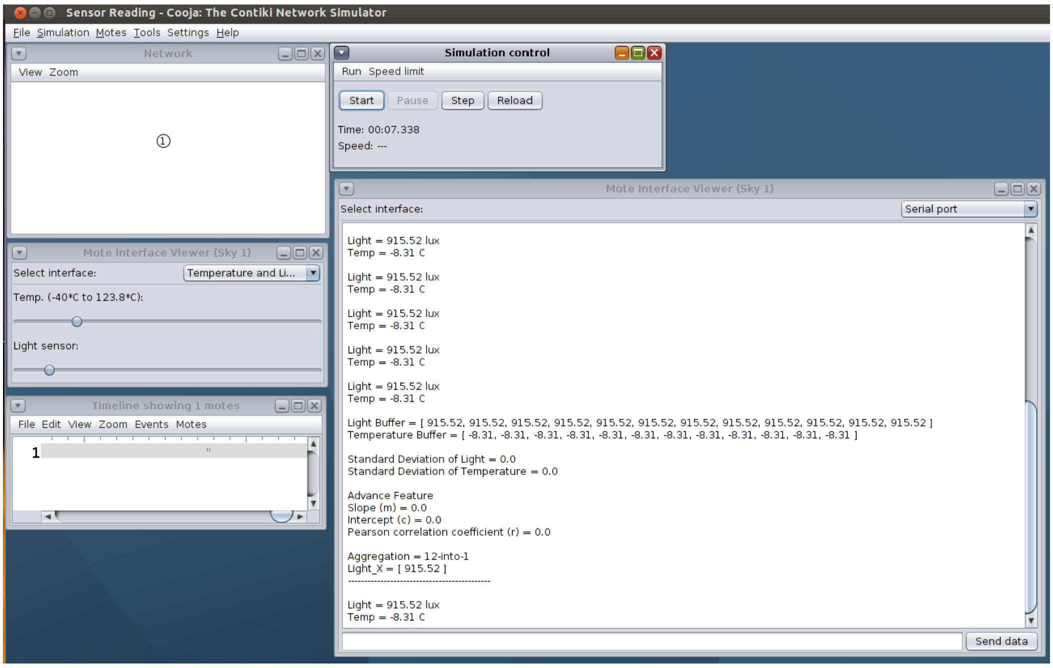

Figure 1: No Relation between Vectors

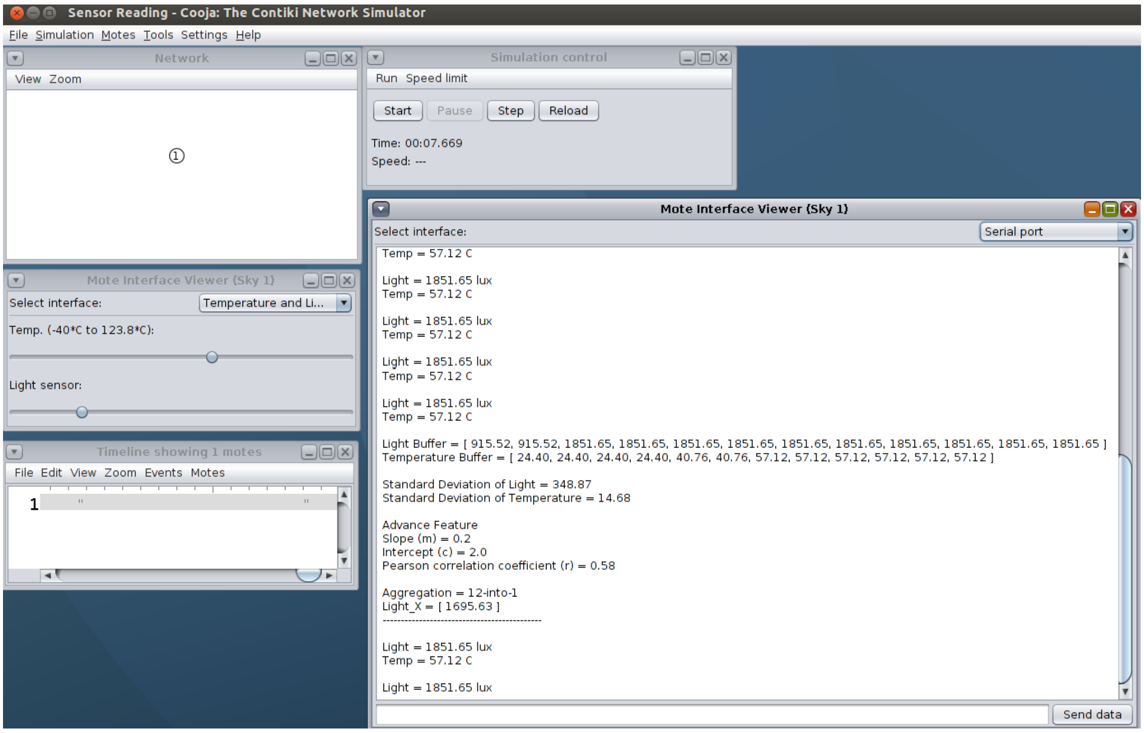

Figure 2: Positive Relation between Vectors