**Error Metric**

Suppose our data set contains K labels $\{Y_1, Y_2, \ldots, Y_m, \ldots, Y_K\}$ that we are predicting.

Let us denote the training dataset labels as $\{Y_1^{train}, Y_2^{train}, \ldots, Y_m^{train}, \ldots, Y_K^{train}\}$. So each $Y_m^{train}$ is a column vector containing entries $Y_{m,i}^{train}$ for each instance $X_i$. For each label, determine the following quantity

$$MaxMin_m := max(Y_m^{train}) - min(Y_m^{train}) = max_i(Y_{m,i}^{train}) - min(Y_{m,i}^{train}).$$

The purpose of this quantity will be to normalize errors for labels that are on different scales.

Now, suppose you have a test dataset of N instances, with given labels $\{Y_1, Y_2, \ldots, Y_m, \ldots, Y_K\}$ and you predict labels $\{\hat{Y}_1, \hat{Y}_2, \ldots, \hat{Y}_m, \ldots, \hat{Y}_K\}$. The error metric is defined as:

$$error = \frac{1}{K}\left(\Sigma_{m=1}^{K}\left(\frac{\text{RMSE of } Y_m}{MaxMin_m}\right)\right) = \frac{1}{K}\left(\Sigma_{m=1}^{K}\left(\frac{\sqrt{\frac{||Y_m - \hat{Y}_m||_2^2}{N}}}{MaxMin_m}\right)\right) = \frac{1}{K}\left(\Sigma_{m=1}^{K}\left(\frac{\sqrt{\frac{\Sigma_{i=1}^{N}\left(|Y_{m,i} - \hat{Y}_{m,i}|^2\right)}{N}}}{MaxMin_m}\right)\right)$$

This error metric will be used to judge your performance.