# MAE and MSE Formula:

## Both mean absolute error (MAE) and mean squared error (MSE) are loss functions used for helping us how is the Accuracy of the predictions across the test set. Their formulas are listed as follows:

Mean Absolute Error Equation:

Mean Squared Error Equation:

# Why someone one use one and not the other?

## From the formulas we can see that the core difference between MAE and MSE is that MSE squares each pair’s error, e.g. if a difference of 2, becomes 4, and this leads to give more weight to outliners under MSE condition.

## Moreover, the optimal prediction of MAE will be their median target value while compared with the optimal prediction of MSE is the mean.

## From simple comparison we can get conclusion that when someone don’t want outliers give a big influence on linear regression, we should use MAE and if we believe that if input data is normally distributed around a mean value, we can use MSE.

# Example situation

For circumstance like sale~customer rate linear regression, price~house squre linear regression, in most circumstance there’s no empty data in data set. However, there may be have several outliner inputs which can have great influence on MSE. So, under this condition we should use MAE. If the data set is checked and is normally distributed, we can use MSE as loss function so can be more appropriate in some cases.

# Run an experiment

In this experiment we used World War II’s climate from Kaggle[1], every day’s Min Temperature in dataset is used as x axis data and every days Max Temperature is used as y axis data. As we can see there are two outliners One is around x axis’s value -20, another is around x axis’s value 10 to 25.

A close up of a map

Description automatically generated

After we use LinearRegression in sklearn library, we got following plot, which is quite fit to graph:

A close up of a map

Description automatically generated

Then we checked loss function, we used MAE got value 3.19932917837853, and used MSE got value 17.631568097568447, and also got RMSE(Root Mean Squared Error) value 4.198996082109204. RMSE, which is a root version of MSE, has very close relationship with MSE. From result we can see that because we have outliners in this dataset so MSE increased greatly and RMSE is around 25% higher than MAE. In this case MAE is more suitable.

**https://peltarion.com/knowledge-center/documentation/evaluation-view/regression-loss-metrics**

<http://zerospectrum.com/2019/06/02/mae-vs-mse-vs-rmse/>

https://www.kaggle.com/smid80/weatherww2