Firstly, both precision and recall are terms used within a classification context in which a model predicts a particular response (in this case, a binary one, such as has disease / does not have disease). Thus, possible outcomes can be false positive, true positive, false negative, and false negative, with the first part of the term referring to whether the model prediction was correct (true) or incorrect (false) and the second part referring to the prediction output by the model (positive or negative). So, in a false positive, for example, the model predicts a positive but was incorrect in doing so. (For example, it predicted the presence of a disease when in reality there was no disease.)

Precision is the number of true positives divided by the sum of true positives and false positives, i.e. TPTP+FP*TP*+*FPTP*​. It is the proportion of the positive classifier predictions that were, in fact, positive. Maximizing this is important since we want our model to be correct when predicting a positive class.

Recall is the number of true positives divided by the sum of true positives and false negatives, i.e. TPTP+FN*TP*+*FNTP*​. Therefore, it is the percentage of total positive cases captured. Recall is important to maximize since we want our model to capture all of the positive cases in the relevant universe.

Both metrics are relevant in evaluating a classification algorithm, but we need to do that carefully, because a change in one leads to a change in the other, and so maximizing both is not an easy task. The tradeoff between the two metrics arises from what aspect of the problem being analyzed is more important in the context of the classification problem at hand.

For example, in the prior example, by not identifying a life-threatening disease in a patient (a false negative), the patient’s life is being put at risk. Likewise with a false positive (deeming someone as having the disease when they do not), the patient may undergo treatment having unpleasant, harmful side effects. Since there may be relative weightings between these outcomes, a tradeoff between precision and recall needs to be considered.