Confusion Matrix

- Let **True positive (TP)** = no. of positive-examples correctly predicted.
- **False negative (FN)** = no. of positive-examples wrongly predicted as negative.
- **False positive (FP)** = no. of negative-examples wrongly predicted as positive.
- **True negative (TN)** = no. of negative-examples correctly predicted.

	Predicted **0**	Predicted **1**
Actual **0**	TN	FP
Actual **1**	FN	TP

Confusion Matrix

- Consider the confusion matrix given below for a binary classifier predicting the presence of a disease
- The classifier made a total of 150 predictions Out of those 150 cases, the classifier predicted "yes" 100 times, and "no" 50 times.
- In reality, 100 patients in the sample have the disease, and 50 patients do not.

	Predicted No	Predicted Yes
Actual No	45	5
Actual Yes	5	95

	Predicted 0	Predicted 1
Actual 0	TN	FP
Actual 1	FN	TP

- **Accuracy:** Overall, how often is the classifier correct?
- $Accuraccy = \frac{TN+TP}{TN+FP+FN+TP}$

$$= \frac{45+95}{150} = 93.33\%$$

	Predicted No	Predicted Yes
Actual No	45	5
Actual Yes	5	95

	Predicted 0	Predicted 1
Actual 0	TN	FP
Actual 1	FN	TP

- **Misclassification Rate:** Overall, how often is it wrong?
- $Missclassification\ Rate = \frac{FN+FP}{TN+FP+FN+TP}$

$$= \frac{5+5}{150} = 6.67\%$$

	Predicted No	Predicted Yes
Actual No	45	5
Actual Yes	5	95

	Predicted 0	Predicted 1
Actual 0	TN	FP
Actual 1	FN	TP

- **True Positive Rate:** When it's actually yes, how often does it predict yes?
- also known as "Sensitivity" or "Recall"
- $$True\ Positive\ rate = \frac{TP}{Actual\ Yes}$$

$$= \frac{95}{100} = 95\%$$

	Predicted No	Predicted Yes
Actual No	45	5
Actual Yes	5	95

	Predicted 0	Predicted 1
Actual 0	TN	FP
Actual 1	FN	TP

- **False Positive Rate:** When it's actually no, how often does it predict yes?

- $False\ Positive\ rate = \frac{FP}{Actual\ No}$

$$= \frac{5}{50} = 10\%$$

	Predicted No	Predicted Yes
Actual No	45	5
Actual Yes	5	95

	Predicted 0	Predicted 1
Actual 0	TN	FP
Actual 1	FN	TP

- **True Negative Rate:** When it's actually no, how often does it predict no?
- also known as "Specificity"

- $$True\ Negative\ rate = \frac{TN}{Actual\ No}$$

$$= \frac{45}{50} = 90\%$$

	Predicted No	Predicted Yes
Actual No	45	5
Actual Yes	5	95

	Predicted 0	Predicted 1
Actual 0	TN	FP
Actual 1	FN	TP

- **Precision:** When it predicts yes, how often is it correct?
- $Precision = \frac{TP}{Predicted\ Yes}$

$$= \frac{95}{100} = 95\%$$

	Predicted No	Predicted Yes
Actual No	45	5
Actual Yes	5	95

	Predicted 0	Predicted 1
Actual 0	TN	FP
Actual 1	FN	TP

- **Prevalence:** How often does the yes condition actually occur in our sample?

- $Prevalence = \frac{Actual\ Yes}{Total}$

$$= \frac{100}{150} = 66.67\%$$

	Predicted No	Predicted Yes
Actual No	45	5
Actual Yes	5	95

	Predicted 0	Predicted 1
Actual 0	TN	FP
Actual 1	FN	TP