Performance measures

Results : confusion matrix

confusion matrix

```
> confusionMatrix(crashTest_1_TEST$LogisPred,crashTest_1_TEST$CarType)
Confusion Matrix and Statistics

           Reference
Prediction  Hatchback SUV
  Hatchback        10   1
  SUV               0   9

               Accuracy : 0.95
                 95% CI : (0.7513, 0.9987)
    No Information Rate : 0.5
    P-Value [Acc > NIR] : 2.003e-05

                  Kappa : 0.9
 Mcnemar's Test P-Value : 1

            Sensitivity : 1.0000
            Specificity : 0.9000
         Pos Pred Value : 0.9091
         Neg Pred Value : 1.0000
             Prevalence : 0.5000
         Detection Rate : 0.5000
   Detection Prevalence : 0.5500
      Balanced Accuracy : 0.9500

       'Positive' Class : Hatchback
```

Confusion matrix

		True condition	
	Total population	Condition positive	Condition negative
Predicted condition	Predicted condition positive	True positive Power Success	False positive Type I error Mistake
	Predicted condition negative	False negative Type II error	True negative Success

Source: https://en.wikipedia.org/wiki/Receiver_operating_characteristic

Measures of performance

- Terminology
  - TP → true positives, TN → true negatives,
  - FP → false positives, FN → false negatives

$$N = TP + TN + FP + FN$$

  - TP – Correct identification of positive labels
  - TN – Correct identification of negative labels
  - FP – Incorrect identification of positive labels
  - FN – Incorrect identification of negative labels

Measures of performance

- Accuracy: Overall effectiveness of a classifier
  - $A = \frac{TP+TN}{N}$
  - Maximum value that accuracy can take is 1
  - This happens when the classifier exactly classifies two groups (i.e, $FP = 0$ and $FN = 0$)
- Remember
  - Total number of true positive labels = TP+FN
- Similarly
  - Total number of true negative labels = TN+FP

TP, TN, FP, FN

Measures of performance

- Sensitivity: Effectiveness of a classifier to identify positive labels
  - $S_e = \frac{TP}{TP + FN}$
- Specificity: Effectiveness of a classifier to identify negative labels
  - $S_p = \frac{TN}{FP + TN}$
- Both S_e and S_p lie between 0 *and* 1, 1 is an ideal value for each of them
- Balanced accuracy
  - $BA = (sensitivity + specificity)/2$

$\boxed{TP}$, TN, FP, FN

Measures of performance

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

$$P = \frac{TP + FN}{N}$$

P

- Positive predictive value: Proportion of correct results in labels identified as positive
    - $PPV = \frac{(sensitivity * prevalence)}{((sensitivity * prevalence) + ((1 - specificity) * (1 - prevalence)))}$
- Negative prediction value: Proportion of correct results in labels identified as negative
    - $NPV = \frac{specificity * (1 - prevalence)}{(((1 - sensitivity) * prevalence) + ((specificity) * (1 - prevalence)))}$

N

Data science for Engineers

Measures of performance

- Detection rate:
    - $DR = \frac{TP}{N}$
- Detection prevalence: prevalence of predicted events
    - $DP = \frac{TP+FP}{N}$
- The Kappa statistic (or value) is a metric that compares an **observed accuracy** with an **expected accuracy** (random chance)
- Kappa = $\frac{observed\ accuracy - expected\ accuracy}{1 - expected\ accuracy}$

Measures of performance

- Observed accuracy
  - $OA = \frac{a+d}{N}$
- Expected accuracy
  - $EA = \frac{(a+c)(a+b)+(b+d)(c+d)}{N}$
- Kappa $= \dfrac{\frac{(a+d)}{N} - \left(\frac{(a+c)(a+b)+(b+d)(c+d)}{N}\right)}{\left(1 - \left(\frac{(a+c)(a+b)+(b+d)(c+d)}{N}\right)\right)}$
  - Where $a, b, c \text{ and } d$ are $TP, FP, FN \text{ and } TN$ respectively

Results : confusion matrix

confusion matrix.

```
> confusionMatrix(crashTest_1_TEST$LogisPred,crashTest_1_TEST$CarType)
Confusion Matrix and Statistics

            Reference
Prediction  Hatchback SUV
  Hatchback        10   1
  SUV               0   9

               Accuracy : 0.95
                 95% CI : (0.7513, 0.9987)
    No Information Rate : 0.5
    P-Value [Acc > NIR] : 2.003e-05

                  Kappa : 0.9
 Mcnemar's Test P-Value : 1

            Sensitivity : 1.0000
            Specificity : 0.9000
         Pos Pred Value : 0.9091
         Neg Pred Value : 1.0000
             Prevalence : 0.5000
         Detection Rate : 0.5000
   Detection Prevalence : 0.5500
      Balanced Accuracy : 0.9500

       'Positive' Class : Hatchback
```

TP = 10 FP = 1

FN = 0 TN = 9

TP FP FN TN

$\frac{19}{20}$

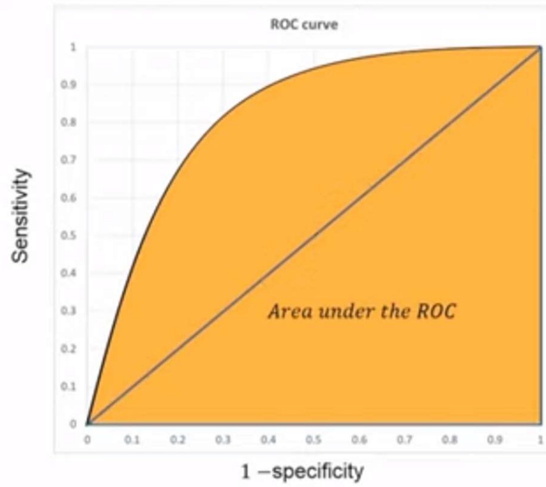
$TP/TP+FN = \frac{10}{10+0}$

$\frac{TP}{TP+FP}$ $\frac{10}{11}$

$\frac{TN}{TN+FN} = \frac{9}{9+0}$

ROC

- ROC can be used to
  - See the classifier performance at different threshold levels (from 0 to 1)
  - AUC- Area under the ROC
    - An area of 1 represents a perfect test; an area of 0.5 represents a worthless model.
    - .90 – 1 = excellent
    - .80 – .90 = good
    - .70 – .80 = fair
    - .60 – .70 = poor
  - AUC < 0.5, check whether your labels are marked in opposite



ROC

- ROC can be used to
  - Compare different classifiers at one threshold or overall threshold levels
  - Performance
  - Model 3 > Model 2 > Model 1