Module 1: Basics and Background

Lecture 3: Error Measures

Prev topic Next topic Prev page
Next page

The Lecture Contains:

- Errors
  - Parameters for errors
  - Example
- ROC curve
- Confusion matrix

Errors

- Positives (P): objects that are "true" answers for a query
- Negatives (N): objects that are not answers for the query
  - $N = DB - P$

Errors

- Positives (P): objects that are "true" answers for a query
- Negatives (N): objects that are not answers for the query
  - $N = DB - P$
- For any querying algorithm,
  - True Positives (TP): Answers that have been found
  - True Negatives (TN): Non-answers that have not been found
  - False Positives (FP): Non-answers that have been found
  - False Negatives (FN): Answers that have not been found
- $P = TP \cup FN$
- $N = TN \cup FP$

Errors

- Positives (P): objects that are "true" answers for a query
- Negatives (N): objects that are not answers for the query
  - $N = DB - P$
- For any querying algorithm,
  - True Positives (TP): Answers that have been found
  - True Negatives (TN): Non-answers that have not been found
  - False Positives (FP): Non-answers that have been found
  - False Negatives (FN): Answers that have not been found
- $P = TP \cup FN$
- $N = TN \cup FP$
- Errors
  - Type I error: FP
  - Type II error: FN

Parameters for errors

- Recall or Sensitivity = $\frac{|TP|}{|TP \cup FN|} = \frac{|TP|}{|P|}$
  - Proportion of answers found

Parameters for errors

- Recall or Sensitivity = $\frac{|TP|}{|TP \cup FN|} = \frac{|TP|}{|P|}$
  - Proportion of answers found
- Precision = $\frac{|TP|}{|TP \cup FP|}$
  - Proportion of "true" answers in those found

Module 1: Basics and Background

Lecture 3: Error Measures

Prev topic Next topic Prev page
Next page

Parameters for errors

- Recall or Sensitivity = $\frac{|TP|}{|TP \cup FN|} = \frac{|TP|}{|P|}$
  - Proportion of answers found
- Precision = $\frac{|TP|}{|TP \cup FP|}$
  - Proportion of "true"answers in those found
- Specificity = $\frac{|TN|}{|TN \cup FP|} = \frac{|TN|}{|N|} = 1 - \frac{|FP|}{|TN \cup FP|} = 1 - \frac{|FP|}{|N|}$
  - Proportion of" "true" non-answers in those not found

Parameters for errors

- Recall or Sensitivity = $\frac{|TP|}{|TP \cup FN|} = \frac{|TP|}{|P|}$
  - Proportion of answers found
- Precision = $\frac{|TP|}{|TP \cup FP|}$
  - Proportion of "true"answers in those found
- Specificity = $\frac{|TN|}{|TN \cup FP|} = \frac{|TN|}{|N|} = 1 - \frac{|FP|}{|TN \cup FP|} = 1 - \frac{|FP|}{|N|}$
  - Proportion of" "true" non-answers in those not found
- F-score or F-measure = $\frac{(1+\beta) \times Precision \times Recall}{\beta \times Precision + Recall}$
  - Single measure capturing both precision and recall
  - Recall is β times more important than precision
  - When β=1, it is the harmonic mean of precision and recall

Parameters for errors

- Recall or Sensitivity = $\frac{|TP|}{|TP \cup FN|} = \frac{|TP|}{|P|}$
  - Proportion of answers found
- Precision = $\frac{|TP|}{|TP \cup FP|}$
  - Proportion of "true" answers in those found
- Specificity = $\frac{|TN|}{|TN \cup FP|} = \frac{|TN|}{|N|} = 1 - \frac{|FP|}{|TN \cup FP|} = 1 - \frac{|FP|}{|N|}$
  - Proportion of" "true" non-answers in those not found
- F-score or F-measure = $\frac{(1+\beta) \times Precision \times Recall}{\beta \times Precision + Recall}$
  - Single measure capturing both precision and recall
  - Recall is β times more important than precision
  - When β=1, it is the harmonic mean of precision and recall
- Accuracy = $\frac{|TP \cup TN|}{|TP \cup TN \cup FP \cup FN|} = \frac{|TP \cup TN|}{|DB|}$
  - Proportion of objects correctly classified as answers and non-answers

Parameters for errors

- Recall or Sensitivity = $\frac{|TP|}{|TP \cup FN|} = \frac{|TP|}{|P|}$
  - Proportion of answers found
- Precision = $\frac{|TP|}{|TP \cup FP|}$
  - Proportion of "true" answers in those found
- Specificity = $\frac{|TN|}{|TN \cup FP|} = \frac{|TN|}{|N|} = 1 - \frac{|FP|}{|TN \cup FP|} = 1 - \frac{|FP|}{|N|}$
  - Proportion of" "true" non-answers in those not found
- F-score or F-measure = $\frac{(1+\beta) \times Precision \times Recall}{\beta \times Precision + Recall}$
  - Single measure capturing both precision and recall
  - Recall is β times more important than precision
  - When β=1, it is the harmonic mean of precision and recall
- Accuracy = $\frac{|TP \cup TN|}{|TP \cup TN \cup FP \cup FN|} = \frac{|TP \cup TN|}{|DB|}$
  - Proportion of objects correctly classified as answers and non-answers
- ROC (Receiver Operating Characteristics) Curve: Sensitivity (y-axis) vs. 1 - Specificity (x-axis)

Module 1: Basics and Background

Lecture 3: Error Measures

Parameters for errors

- Recall or Sensitivity = $\frac{|TP|}{|TP \cup FN|} = \frac{|TP|}{|P|}$
  - Proportion of answers found
- Precision = $\frac{|TP|}{|TP \cup FP|}$
  - Proportion of "true" answers in those found
- Specificity = $\frac{|TN|}{|TN \cup FP|} = \frac{|TN|}{|N|} = 1 - \frac{|FP|}{|TN \cup FP|} = 1 - \frac{|FP|}{|N|}$
  - Proportion of" "true" non-answers in those not found
- F-score or F-measure = $\frac{(1+\beta) \times Precision \times Recall}{\beta \times Precision + Recall}$
  - Single measure capturing both precision and recall
  - Recall is β times more important than precision
  - When β=1, it is the harmonic mean of precision and recall
- Accuracy = $\frac{|TP \cup TN|}{|TP \cup TN \cup FP \cup FN|} = \frac{|TP \cup TN|}{|DB|}$
  - Proportion of objects correctly classified as answers and non-answers
- ROC (Receiver Operating Characteristics) Curve: Sensitivity (y-axis) vs. 1 - Specificity (x-axis)
- Confusion matrix: Found out by algorithm (predictions) on rows vs "true" answers (actuals) on columns

Example

$DB = \{O_1, O_2, O_3, O_4, O_5, O_6, O_7, O_8\}$

Correct answer set = $\{O_1, O_5, O_7\}$

Algorithm returns = $\{O_1, O_3, O_5, O_6\}$

Example

$DB = \{O_1, O_2, O_3, O_4, O_5, O_6, O_7, O_8\}$

Correct answer set = $\{O_1, O_5, O_7\}$

Algorithm returns = $\{O_1, O_3, O_5, O_6\}$

$\therefore \quad P =$

$N =$

$TP =$

$TN =$

$FP =$

$FN =$

Example

$DB = \{O_1, O_2, O_3, O_4, O_5, O_6, O_7, O_8\}$

Correct answer set = $\{O_1, O_5, O_7\}$

Algorithm returns = $\{O_1, O_3, O_5, O_6\}$

$\therefore \quad P = \{O_1, O_5, O_7\}$

$N =$

$TP =$

$TN =$

$FP =$

$FN =$

Example

$DB = \{O_1, O_2, O_3, O_4, O_5, O_6, O_7, O_8\}$

Correct answer set = $\{O_1, O_5, O_7\}$

Algorithm returns = $\{O_1, O_3, O_5, O_6\}$

$\therefore \quad P = \{O_1, O_5, O_7\}$

$N = \{O_2, O_3, O_4, O_6, O_8\}$

$TP =$

$TN =$

$FP =$

$FN =$

Example

$DB = \{O_1, O_2, O_3, O_4, O_5, O_6, O_7, O_8\}$

Correct answer set = $\{O_1, O_5, O_7\}$

Algorithm returns = $\{O_1, O_3, O_5, O_6\}$

$\therefore \quad P = \{O_1, O_5, O_7\}$

$N = \{O_2, O_3, O_4, O_6, O_8\}$

$TP = \{O_1, O_5\}$

$TN =$

$FP =$

$FN =$

Example

$DB = \{O_1, O_2, O_3, O_4, O_5, O_6, O_7, O_8\}$

Correct answer set = $\{O_1, O_5, O_7\}$

Algorithm returns = $\{O_1, O_3, O_5, O_6\}$

$\therefore \quad P = \{O_1, O_5, O_7\}$

$N = \{O_2, O_3, O_4, O_6, O_8\}$

$TP = \{O_1, O_5\}$

$TN = \{O_2, O_4, O_8\}$

$FP =$

$FN =$

Example

$DB = \{O_1, O_2, O_3, O_4, O_5, O_6, O_7, O_8\}$

Correct answer set = $\{O_1, O_5, O_7\}$

Algorithm returns = $\{O_1, O_3, O_5, O_6\}$

$\therefore \quad P = \{O_1, O_5, O_7\}$

$N = \{O_2, O_3, O_4, O_6, O_8\}$

$TP = \{O_1, O_5\}$

$TN = \{O_2, O_4, O_8\}$

$FP = \{O_3, O_6\}$

$FN =$

Example

$DB = \{O_1, O_2, O_3, O_4, O_5, O_6, O_7, O_8\}$

Correct answer set = $\{O_1, O_5, O_7\}$

Algorithm returns = $\{O_1, O_3, O_5, O_6\}$

$\therefore \quad P = \{O_1, O_5, O_7\}$

$N = \{O_2, O_3, O_4, O_6, O_8\}$

$TP = \{O_1, O_5\}$

$TN = \{O_2, O_4, O_8\}$

$FP = \{O_3, O_6\}$

$FN = \{O_7\}$

Example

$DB = \{O_1, O_2, O_3, O_4, O_5, O_6, O_7, O_8\}$

Correct answer set = $\{O_1, O_5, O_7\}$

Algorithm returns = $\{O_1, O_3, O_5, O_6\}$

$\therefore \quad P = \{O_1, O_5, O_7\}$

$N = \{O_2, O_3, O_4, O_6, O_8\}$

$TP = \{O_1, O_5\}$

$TN = \{O_2, O_4, O_8\}$

$FP = \{O_3, O_6\}$

$FN = \{O_7\}$

$\therefore$ Recall = Sensitivity =

Precision =

Specificity =

F-score =

Accuracy =

Example

$DB = \{O_1, O_2, O_3, O_4, O_5, O_6, O_7, O_8\}$

Correct answer set = $\{O_1, O_5, O_7\}$

Algorithm returns = $\{O_1, O_3, O_5, O_6\}$

$\therefore \quad P = \{O_1, O_5, O_7\}$

$N = \{O_2, O_3, O_4, O_6, O_8\}$

$TP = \{O_1, O_5\}$

$TN = \{O_2, O_4, O_8\}$

$FP = \{O_3, O_6\}$

$FN = \{O_7\}$

$\therefore$ Recall = Sensitivity = 2/3 = 0.67

Precision =

Specificity =

F-score =

Accuracy =

Example

$DB = \{O_1, O_2, O_3, O_4, O_5, O_6, O_7, O_8\}$

Correct answer set = $\{O_1, O_5, O_7\}$

Algorithm returns = $\{O_1, O_3, O_5, O_6\}$

$\therefore \quad P = \{O_1, O_5, O_7\}$

$N = \{O_2, O_3, O_4, O_6, O_8\}$

$TP = \{O_1, O_5\}$

$TN = \{O_2, O_4, O_8\}$

$FP = \{O_3, O_6\}$

$FN = \{O_7\}$

$\therefore$ Recall = Sensitivity = 2/3 = 0.67

Precision = 2/4 = 0.5

Specificity =

F-score =

Accuracy =

Example

$DB = \{O_1, O_2, O_3, O_4, O_5, O_6, O_7, O_8\}$

Correct answer set = $\{O_1, O_5, O_7\}$

Algorithm returns = $\{O_1, O_3, O_5, O_6\}$

$\therefore \quad P = \{O_1, O_5, O_7\}$

$N = \{O_2, O_3, O_4, O_6, O_8\}$

$TP = \{O_1, O_5\}$

$TN = \{O_2, O_4, O_8\}$

$FP = \{O_3, O_6\}$

$FN = \{O_7\}$

$\therefore$ Recall = Sensitivity = 2/3 = 0.67

Precision = 2/4 = 0.5

Specificity = 3/5 = 0.6

F-score =

Accuracy =

Example

$DB = \{O_1, O_2, O_3, O_4, O_5, O_6, O_7, O_8\}$

Correct answer set = $\{O_1, O_5, O_7\}$

Algorithm returns = $\{O_1, O_3, O_5, O_6\}$

$\therefore \quad P = \{O_1, O_5, O_7\}$

$N = \{O_2, O_3, O_4, O_6, O_8\}$

$TP = \{O_1, O_5\}$

$TN = \{O_2, O_4, O_8\}$

$FP = \{O_3, O_6\}$

$FN = \{O_7\}$

$\therefore$ Recall = Sensitivity = 2/3 = 0.67

Precision = 2/4 = 0.5

Specificity = 3/5 = 0.6

F-score = 4/7 = 0.571

Accuracy =

Example

$D = \{O_1, O_2, O_3, O_4, O_5, O_6, O_7, O_8\}$

Correct answer set = $\{O_1, O_5, O_7\}$

Algorithm returns = $\{O_1, O_3, O_5, O_6\}$

$\therefore \quad P = \{O_1, O_5, O_7\}$

$N = \{O_2, O_3, O_4, O_6, O_8\}$

$TP = \{O_1, O_5\}$

$TN = \{O_2, O_4, O_8\}$

$FP = \{O_3, O_6\}$

$FN = \{O_7\}$

$\therefore$ Recall = Sensitivity = 2/3 = 0.67

Precision = 2/4 = 0.5

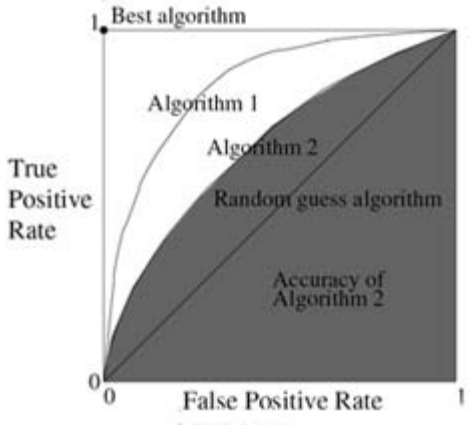
Specificity = 3/5 = 0.6

F-score = 4/7 = 0.571

Accuracy = 5/8 = 0.625

ROC curve

- Sensitivity vs 1 - specificity
- True positive rate vs false positive rate
- A random guess algorithm is a 45^0 line
- Area under the curve measures accuracy (or "discrimination")
- A perfect algorithm has an area of 1
- A random (or useless) algorithm has an area of 0.5
- Grading
- 0.9-1.0: excellent; 0.8-0.9: good; 0.7-0.8: fair; 0.6-0.7: poor; 0.5-0.6: fail



Confusion matrix

- Used more in data mining methods
- Shows which classes are harder to identify

		True answers	
		Positives P	Negatives N
Found by algorithm	Positives P'	TP	FP
	Negatives N'	FN	TN