How Accurate Is It? Introducing the Difference/Error Matrix

Purpose

To quantitatively evaluate the accuracy of a classification

Overview

Students will sort birds into three possible classes: carnivores (meat eaters), herbivores (plant eaters), and omnivores (meat and plant eaters) based on the bird's beak. They will then compare their answers with a given set of validation data and generate a difference/error matrix. The students will then discuss how to improve their accuracy based on identifying specific mistakes they made as indicated by the difference/error matrix.

Time

One class period

Level

Intermediate to Advanced

Prerequisites

Basic ability to classify things

Fractions and percentages

Key Concepts

- Classification helps us organize and understand the natural world.
- In order for classification systems to be useful, we need to quantitatively determine their accuracy.
- Criteria are used to define accuracy levels.

Skills

- *Classifying* birds
- *Evaluating* the accuracy of the classification
- *Improving* the accuracy of the classification based on the evaluation
- *Analyzing* data to understand the inter-relationships of a classification and its accuracy
- *Identifying* decision criteria for a classification system
- *Collecting* and *interpreting* validation data
- *Building* and *analyzing* a difference/error matrix for accuracy assessment
- *Solving* problems cooperatively to resolve accuracy issues

Materials and Tools

- Master set of bird pictures
- Master validation sheet
- Overhead showing a sample bird classification work sheet
- Set of bird pictures
- Sample beak sketches
- Classification Work Sheet
- Difference/Error Matrix Work Sheet.

Preparation

Bird picture sets need to be reproduced without the answers on the back. Also student work sheets need to be reproduced for each group.

Background

Scientists classify many features of our environment, such as species of life, forest types, or soil types. These classifications are a fundamental mechanism for helping us to organize and to understand the natural world. There may be several different appropriate ways to classify a set of objects of interest. Two particular objects may be classified differently either because of error on the part of one or both of the classifiers, or simply because different classifying criteria were used. In any case, we need to know how much error is in our classification in order to use the information we have obtained with some confidence in its accuracy. Ultimately, the information generated by the classification of remotely sensed data will be used to make important decisions about global problems such as deforestation, global warming, and environmental degradation. It is very important that we not make these decisions based on information that is inaccurate.

A difference/error matrix is the basic tool used for accuracy assessment of remotely sensed data. It gives us a mechanism for generating a number rating the overall accuracy of a classification or map and provides information about the sources of error. This can focus our attention on those areas or classes that require it. We can use this information to improve the quality of our classification criteria, and to improve our skill at distinguishing those classes for which there is a lot of confusion.

References

Peterson's Field Guide to Birds

Audubon Field Guides

The Illustrated Encyclopedia of Birds: The Definitive Reference to Birds of the World. Consultant-in Chief Dr. C. Perrins. New York: Prentice Hall Press, 1990.

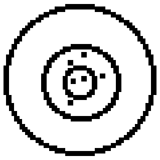
Check local resources for regional guides

Acknowledgment

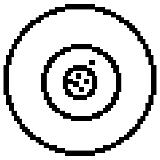
Art by Linda Isaacson

Key Terms and Concepts

accuracy: the degree of conformity to a standard or accepted value. Compare to precision.



The marks on this bull's-eye have high accuracy and low precision



The marks on this bull's-eye have high accuracy and high precision

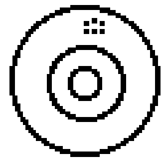
classification: taking a set or group of items and sorting them (classifying them) into well-defined and distinct subsets according to specific criteria. For example, taking a map and outlining areas of evergreen trees, deciduous trees, mixed evergreen and deciduous trees, and non-forest.

criteria: a decision rule. For example, if a forest stand has more than 50% evergreen needles in its canopy, the stand will be classified as evergreen. The preceding definition (e.g., more than 50% evergreen needles) is the *criteria*, the *category* or *class* is evergreen.

dataset: a group of values related to the same question being asked. These values will be analyzed together as a group. For example, the set of the heights of all students in this class would be one dataset.

difference/error matrix: (see the difference/error matrix on the work sheet at the end of this exercise) a table of numbers organized in rows and columns which compares a classification to validation data. The columns represent the validation data while the rows represent the classification generated by students. A difference/error matrix is a very effective way to represent accuracy. Correct and incorrect classifications can be compared for each category and used to improve the accuracy of the original classification.

precision: the closeness of several measures to each other. The repeatability of a measurement. This is a very important part of any scientific operation, but is different from accuracy.



The marks on this bull's-eye have high precision and low accuracy

validation data: data collected with a presumed high degree of accuracy. A classification of items (birds in this exercise) is compared to validation data: 1.) to improve the decision criteria for the classification 2.) to better understand the sources of error in the classification; and 3.) to assess the accuracy of the classification data.

Validation data is often collected to improve the classification of an image generated by some form of remote sensing (aerial photography or satellite imagery). Often the term "ground truth" is used in place of validation data, however, many scientists prefer the term reference or validation data. Data that is gathered on the ground always has some degree of error and thus does not represent the "truth".

Example

The following is an example of a filled in classification work sheet, difference/error matrix, and an overall accuracy calculation.

Table LAND-L-1: Sample Bird Classification Work Sheet

Bird Id#	Student Classification	Validation Data	✔ or x
1	Carnivore	Carnivore	✔
2	Omnivore	Carnivore	x
3	Herbivore	Herbivore	✔
4	Carnivore	Carnivore	✔
5	Herbivore	Herbivore	✔
6	Herbivore	Omnivore	x
7	Omnivore	Omnivore	✔
8	Carnivore	Carnivore	✔
9	Carnivore	Herbivore	x
10	Omnivore	Carnivore	x

Table LAND-L-2: Sample Difference/Error Matrix Work Sheet

Validation Data

Student Data	Carnivore	Herbivore	Omnivore	Row Total
Carnivore	A1. 3	B1. 1	C1. 0	D1. 4
Herbivore	A2. 0	B2. 2	C2. 1	D2. 3
Omnivore	A3. 2	B3. 0	C3. 1	D3. 3
Column Total	A4. 5	B4. 3	C4. 2	D4. 10

Note: Row and column totals should add up to the same number. Check with others in your group to make sure you counted correctly for each answer in the matrix.

$$D4 = (A4 + B4 + C4) = (D1 + D2 + D3)$$

(column total) (row total)

How to read this information:

Across row one (A1-D1) of this example, three carnivores were correctly identified by the students as carnivores, one herbivore was incorrectly classified as a carnivore and no omnivores were classified as carnivores.

Computing the Accuracy:

$$\text{Overall accuracy} = \frac{\text{sum of major diagonal (A1+B2+C3)}}{\text{total of entire matrix (D4)}}$$

Step 1: Sum the values in the boxes along the major diagonal (A1+B2+C3) shown in Table 4-13: Sample Difference/Error Matrix. This number is the total number of correct classifications. In this example there are six correct classifications out of ten total samples.

(3+2+1) = 6

Step 2: Divide the total number of correct classifications (A1+B2+C3) by the total number of samples (box D4).

6 divided by 10 = 0.6

Step 3: Multiply by 100 for the overall accuracy of the exercise:

0.6 times 100 = 60% accuracy

The calculation can be done for any of the individual categories as well (e.g., 3 out of 5 carnivores were classified correctly). The numbers off the major diagonal represent "incorrect" classifications. Each error or difference is an omission from the correct category and a commission (i.e., an erroneous addition) to the incorrect category.

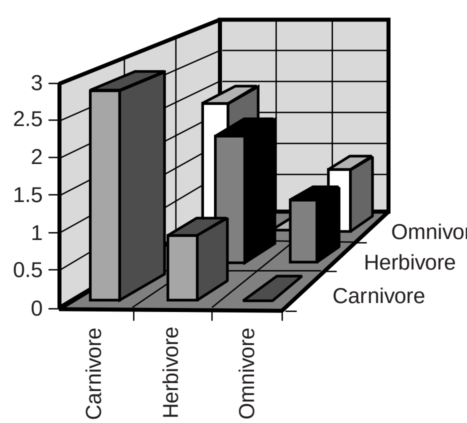
If your answer is between:	Your Level of Expertise is:
0%-50%	Novice
51%-85%	Intermediate
86%-100%	Advanced

The class can also compare fractions (1/2 is less than 3/4, 3/4 is less than 9/10) instead of percentages.

Adaptations

1. A visual interpretation can be used instead of mathematically calculating the overall accuracy. Layout a 3 cell x 3 cell grid on a sheet of paper numbered like the cells in the difference/error matrix.
   Visually represent the number of birds in each box by either graphing or physically stacking blocks in the boxes. The tallest columns should be along the diagonal of the grid.
2. If the class has access to computer spreadsheets, a 3-D graph can be created to represent the answers. Figure LAND-L-2 shows the data from the example difference/error matrix graphed in a
   3-D format.
3. The activity may be modified by leading the activity for the whole group and creating one difference/error matrix on the black board.

Figure LAND-L-2: Difference/Error Matrix of Bird Classification Data



Source: GLOBE, 1996

What To Do and How To Do It

1. To prepare your students, discuss with them the following questions:
   - Why do we organize or sort objects into groups?
   - How do we sort these objects?
   - Name three examples of objects that are commonly sorted into groups.
2. Copy and distribute the student work sheets, the bird pictures, the bird beak sketches, the classification work sheet, and the difference error/matrix work sheet.
3. Have your students follow the instructions on the work sheets, to do the following steps:
   - Classifying pictures of birds into three categories.
   - Comparing answers with the reference data provided.
   - Generating a difference/error matrix using the results of the comparison.
4. After your students have completed this activity, discuss the results with your students by asking the following questions:
   - How did different students' results vary?
   - Why do students think this happened?
   - What other classifications might be compared using a difference/error matrix (e.g., maps identifying land cover for a specific location versus carefully checking the same location in person.