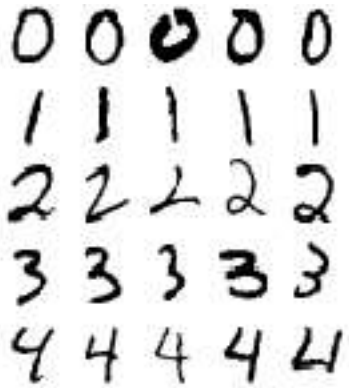
Confusion Matrix

- If n classes, n x n matrix comparing actual versus predicted classes.

Actual Class	Class Predicted by Model: Cat	Class Predicted by Model: Dog
Cat	9	1
Dog	4	6

Example: Handwritten Digit Recognition

Class	1	2	3	4	5	6	7	8	9	0	Error Rate
1	191	0	0	5	1	0	1	0	2	0	4.5
2	0	188	2	0	0	1	3	0	0	6	6.0
3	0	3	191	0	1	0	2	0	3	0	4.5
4	8	0	0	187	4	0	1	0	0	0	6.5
5	0	0	0	0	193	0	0	0	7	0	3.5
6	0	0	0	0	1	196	0	2	0	1	2.0
7	2	2	0	2	0	1	190	0	1	2	5.0
8	0	1	0	0	1	2	2	196	0	0	2.0
9	5	0	2	0	8	0	3	0	179	3	10.5
0	1	4	0	0	0	1	1	0	1	192	4.5
									Overall error rate		4.85%



Slide from Nelson Morgan at ICSI / Berkeley

Example: Handwritten Digit Recognition

- Visualization from MathWorks

Confusion Matrix

- For binary classifier, can call one class positive, the other negative.
- Should we call cats positive or negative?

Photo from [1]

Actual Class	Class Predicted by Model: Cat	Class Predicted by Model: Dog
Cat	9	1
Dog	4	6

Confusion Matrix

- For binary classifier, can call one class positive, the other negative.
- Cats are cuter, so cats = positive.

		Class Predicted by Model	
		Positive	**Negative**
Actual Class	**Positive**	9	1
	Negative	4	6

Confusion Matrix

- For binary classifier, can call one class positive, the other negative.

		Class Predicted by Model	
		Positive	**Negative**
Actual Class	**Positive**	True positive	1
	Negative	4	True negative

Confusion Matrix

- For binary classifier, can call one class positive, the other negative.

		Class Predicted by Model	
		Positive	**Negative**
Actual Class	**Positive**	True positive	**?**
	Negative	?	True negative

Confusion Matrix

- For binary classifier, can call one class positive, the other negative.

		Class Predicted by Model	
		Positive	**Negative**
Actual Class	**Positive**	True positive	False negative
	Negative	False positive	True negative

Classifier Performance

- **Accuracy**: (TP + TN) / (Population Size)
- **Precision**: TP / (Predicted Positives) = (TP + FP)
- **Recall**: TP / (Actual Positives) = TP / (TP + FN)
  (also known as **sensitivity, true positive rate**)
- **Specificity**: TN / (Actual Negatives) = TN / (TN + FP)
  (also known as **true negative rate**)

		Class Predicted by Model	
		Positive	**Negative**
Actual Class	**Positive**	True positive	False negative
	Negative	False positive	True negative

Classifier Performance: Unbalanced Dataset

- **Accuracy**: (TP + TN) / (Population Size)
- **Precision**: TP / (Predicted Positives) = (TP + FP)
- **Recall**: TP / (Actual Positives) = TP / (TP + FN)
  (also known as **sensitivity, true positive rate**)
- **Specificity**: TN / (Actual Negatives) = TN / (TN + FP)
  (also known as **true negative rate**)

		Class Predicted by Model	
		Positive	**Negative**
Actual Class	**Positive**	TP: 99	FN: 0
	Negative	FP: 1	TN: 0

Classifier Performance: Unbalanced Dataset

- **Accuracy**: (TP + TN) / (Population Size) = ?
- **Precision**: TP / (Predicted Positives) = (TP + FP)
- **Recall**: TP / (Actual Positives) = TP / (TP + FN)
  (also known as **sensitivity, true positive rate**)
- **Specificity**: TN / (Actual Negatives) = TN / (TN + FP)
  (also known as **true negative rate**)

		Class Predicted by Model	
		Positive	**Negative**
Actual Class	**Positive**	TP: 99	FN: 0
	Negative	FP: 1	TN: 0

Classifier Performance: Unbalanced Dataset

- **Accuracy**: (TP + TN) / (Population Size) = 99%
- **Precision**: TP / (Predicted Positives) = (TP + FP) = ?
- **Recall**: TP / (Actual Positives) = TP / (TP + FN)
  (also known as **sensitivity, true positive rate**)
- **Specificity**: TN / (Actual Negatives) = TN / (TN + FP)
  (also known as **true negative rate**)

		Class Predicted by Model	
		Positive	**Negative**
Actual Class	**Positive**	TP: 99	FN: 0
	Negative	FP: 1	TN: 0

Classifier Performance: Unbalanced Dataset

- **Accuracy**: (TP + TN) / (Population Size) = 99%
- **Precision**: TP / (Predicted Positives) = (TP + FP) = 99%
- **Recall**: TP / (Actual Positives) = TP / (TP + FN)
  (also known as **sensitivity, true positive rate**)
- **Specificity**: TN / (Actual Negatives) = TN / (TN + FP)
  (also known as **true negative rate**)

		Class Predicted by Model	
		Positive	**Negative**
Actual Class	**Positive**	TP: 99	FN: 0
	Negative	FP: 1	TN: 0

Classifier Performance: Unbalanced Dataset

- **Accuracy**: (TP + TN) / (Population Size) = 99%
- **Precision**: TP / (Predicted Positives) = (TP + FP) = 99%
- **Recall**: TP / (Actual Positives) = TP / (TP + FN) = ?
  (also known as **sensitivity, true positive rate**)
- **Specificity**: TN / (Actual Negatives) = TN / (TN + FP)
  (also known as **true negative rate**)

		Class Predicted by Model	
		Positive	**Negative**
Actual Class	**Positive**	TP: 99	FN: 0
	Negative	FP: 1	TN: 0

Classifier Performance: Unbalanced Dataset

- **Accuracy**: (TP + TN) / (Population Size) = 99%
- **Precision**: TP / (Predicted Positives) = (TP + FP) = 99%
- **Recall**: TP / (Actual Positives) = TP / (TP + FN) = 100%
  (also known as **sensitivity, true positive rate**)
- **Specificity**: TN / (Actual Negatives) = TN / (TN + FP)
  (also known as **true negative rate**)

		Class Predicted by Model	
		Positive	**Negative**
Actual Class	**Positive**	TP: 99	FN: 0
	Negative	FP: 1	TN: 0

Classifier Performance: Unbalanced Dataset

- **Accuracy**: (TP + TN) / (Population Size) = 99%
- **Precision**: TP / (Predicted Positives) = (TP + FP) = 99%
- **Recall**: TP / (Actual Positives) = TP / (TP + FN) = 100%
  (also known as **sensitivity, true positive rate**)
- **Specificity**: TN / (Actual Negatives) = 0%
  (also known as **true negative rate**)

		Class Predicted by Model	
		Positive	**Negative**
Actual Class	**Positive**	TP: 99	FN: 0
	Negative	FP: 1	TN: 0