Machine Learning

Classification

Curtis Larsen

Utah Tech University—Computing

Spring 2025

Objectives

Objectives:

- Understand classification tasks
- Identify classification problems
- Review common classification algorithms
- Determine effectiveness of classification models
- Train models using classification algorithms
- Select best models

What is classification?

Definition

- Binary classifier
- Multi-class
- One vs all (OVA) or One-vs-rest
- One vs one

Example Task

Task Identification

Which are classification?

- Given an image of a cell, predict if it is cancerous.

Task Identification

Which are classification?

- Given an image of a cell, predict if it is cancerous.
- Given an image of a cell, predict its size.

Task Identification

Which are classification?

- Given an image of a cell, predict if it is cancerous.
- Given an image of a cell, predict its size.
- Given an image of a digit, predict its numerical value.

Task Identification

Which are classification?

- Given an image of a cell, predict if it is cancerous.
- Given an image of a cell, predict its size.
- Given an image of a digit, predict its numerical value.
- Given demographic information of a passenger on the titanic, predict if they will survive.

Task Identification

Which are classification?

- Given an image of a cell, predict if it is cancerous.
- Given an image of a cell, predict its size.
- Given an image of a digit, predict its numerical value.
- Given demographic information of a passenger on the titanic, predict if they will survive.
- Given transcript information of a new graduate, predict their starting salary.

Task Identification

Which are classification?

- Given an image of a cell, predict if it is cancerous.
- Given an image of a cell, predict its size.
- Given an image of a digit, predict its numerical value.
- Given demographic information of a passenger on the titanic, predict if they will survive.
- Given transcript information of a new graduate, predict their starting salary.
- Given transcript information of a new graduate, predict their starting tax bracket.

Task Identification

Which are classification?

- Given an image of a cell, predict if it is cancerous.
- Given an image of a cell, predict its size.
- Given an image of a digit, predict its numerical value.
- Given demographic information of a passenger on the titanic, predict if they will survive.
- Given transcript information of a new graduate, predict their starting salary.
- Given transcript information of a new graduate, predict their starting tax bracket.
- Given medical and demographic information of a patient, predict their probability of a heart attack in the next 24 months.

Task Identification

Which are classification?

- Given an image of a cell, predict if it is cancerous.
- Given an image of a cell, predict its size.
- Given an image of a digit, predict its numerical value.
- Given demographic information of a passenger on the titanic, predict if they will survive.
- Given transcript information of a new graduate, predict their starting salary.
- Given transcript information of a new graduate, predict their starting tax bracket.
- Given medical and demographic information of a patient, predict their probability of a heart attack in the next 24 months.

Example Data

Common Algorithms

List

- Decision Tree
- Logistic Regression (Yes, it's a classifier)
- Stochastic Gradient Descent
- Nearest Neighbors
- Support Vector Classification
- Random Forest
- Gradient Boosting

Metrics

Sample Numbers

There are 60,000 samples for a binary classification task.

5,421 of the samples are the positive case.

54,579 of the samples are the negative case.

A certain binary classifier predicts 3,530 of the positive samples correctly, and 53,892 of the negative samples correctly.

Accuracy

Percentage of predictions that were correct.

Accuracy

Percentage of predictions that were correct.

In our sample case

Accuracy $= (53,892 + 3,530)/60,000 = 0.957$.

Accuracy

Percentage of predictions that were correct.

In our sample case

Accuracy $= (53,892 + 3,530)/60,000 = 0.957$.

Sounds pretty good.

Confusion Matrix

	Predicted Negative	Predicted Positive
Actual Negative	True-Negative	False-Positive
Actual Positive	False-Negative	True-Positive

Confusion Matrix

	Predicted Negative	Predicted Positive
Actual Negative	True-Negative	False-Positive
Actual Positive	False-Negative	True-Positive

In our sample case

	Predicted Negative	Predicted Positive
Actual Negative	53,892	687
Actual Positive	1,891	3,530

Confusion Matrix

	Predicted Negative	Predicted Positive
Actual Negative	True-Negative	False-Positive
Actual Positive	False-Negative	True-Positive

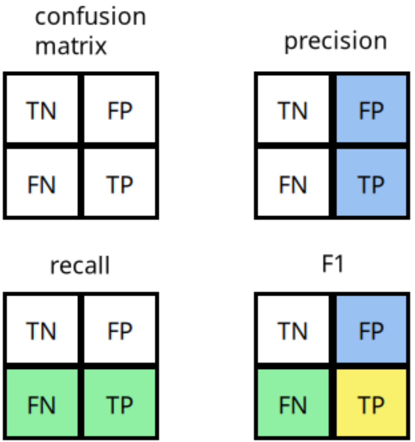
In our sample case

	Predicted Negative	Predicted Positive
Actual Negative	53,892	687
Actual Positive	1,891	3,530

Do we still feel good about the quality?

Confusion Matrix - Metrics

- TN, FP, FN, TP
- precision $= \frac{TP}{TP+FP}$
- recall $= \frac{TP}{TP+FN}$
- $F_1 = \frac{2}{\frac{1}{\text{precision}} + \frac{1}{\text{recall}}}$
- $F_1 = \frac{TP}{TP + \frac{FP+FN}{2}}$



Precision-Recall Example

In our sample case

	Predicted Negative	Predicted Positive
Actual Negative	53,892	687
Actual Positive	1,891	3,530

Precision-Recall Example

In our sample case

	Predicted Negative	Predicted Positive
Actual Negative	53,892	687
Actual Positive	1,891	3,530

precision $=$

Precision-Recall Example

In our sample case

	Predicted Negative	Predicted Positive
Actual Negative	53,892	687
Actual Positive	1,891	3,530

precision $= 3530/(687 + 3530) = 0.837$

Precision-Recall Example

In our sample case

	Predicted Negative	Predicted Positive
Actual Negative	53,892	687
Actual Positive	1,891	3,530

precision $= 3530/(687 + 3530) = 0.837$

recall $=$

Precision-Recall Example

In our sample case

	Predicted Negative	Predicted Positive
Actual Negative	53,892	687
Actual Positive	1,891	3,530

precision $=3530/(687 + 3530) = 0.837$

recall $=3530/(1891 + 3530) = 0.651$

Precision-Recall Example

In our sample case

	Predicted Negative	Predicted Positive
Actual Negative	53,892	687
Actual Positive	1,891	3,530

precision $= 3530/(687 + 3530) = 0.837$

recall $= 3530/(1891 + 3530) = 0.651$

$F_1 =$

Precision-Recall Example

In our sample case

	Predicted Negative	Predicted Positive
Actual Negative	53,892	687
Actual Positive	1,891	3,530

precision $=3530/(687+3530)=0.837$

recall $=3530/(1891+3530)=0.651$

$F_1 = 3530/(3530+((687+1891)/2)) = 0.732$

Precision-Recall Example

In our sample case

	Predicted Negative	Predicted Positive
Actual Negative	53,892	687
Actual Positive	1,891	3,530

precision $=3530/(687+3530)=0.837$

recall $=3530/(1891+3530)=0.651$

$F_1 = 3530/(3530+((687+1891)/2))=0.732$

Precision-Recall Example

In our sample case

	Predicted Negative	Predicted Positive
Actual Negative	53,892	687
Actual Positive	1,891	3,530

precision $= 3530/(687 + 3530) = 0.837$

recall $= 3530/(1891 + 3530) = 0.651$

$F_1 = 3530/(3530 + ((687 + 1891)/2)) = 0.732$

Now how do we feel about the quality?

Limits

Metric	minimum	maximum
precision $= \frac{TP}{TP+FP}$		

Limits

Metric	minimum	maximum
precision $= \frac{TP}{TP+FP}$	0.0	1.0
recall $= \frac{TP}{TP+FN}$		

Limits

Metric	minimum	maximum
precision $= \frac{TP}{TP+FP}$	0.0	1.0
recall $= \frac{TP}{TP+FN}$	0.0	1.0
$F_1 = \frac{TP}{TP+\frac{FP+FN}{2}}$		

Limits

Metric	minimum	maximum
precision $= \frac{TP}{TP+FP}$	0.0	1.0
recall $= \frac{TP}{TP+FN}$	0.0	1.0
$F_1 = \frac{TP}{TP+\frac{FP+FN}{2}}$	0.0	1.0

Summary

- Accuracy
- Confusion Matrix
- Precision
- Recall
- F_1

Implementation

Summary

-
-
-