# CSE 4/587
## Data Intensive Computing

Dr. Eric Mikida
epmikida@buffalo.edu
208 Capen Hall

Dr. Shamshad Parvin
shamsadp@buffalo.edu
313 Davis Hall

# Evaluating Classifiers

# Announcements

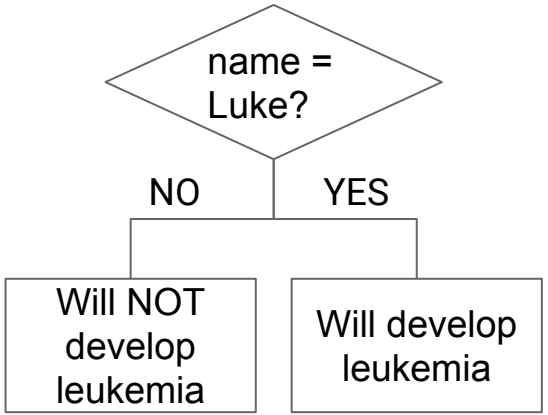
- Project Phase 2 released for those that want to start early
  - Must complete Phase 1 first
- Midterm next Wednesday
  - Practice exam will be posted shortly
  - Monday will be a review class
  - Exam will be held in Cooke 121 **AND** Alumni 97, keep an eye on Piazza

# Evaluating Classifiers

- Classifiers automatically label data as belonging to one class or another
- How can we determine the effectiveness of our models?
- Accuracy is sometimes a reasonable first cut…but not always a wise choice…

# Pitfalls with using Accuracy

- Take the following decision tree for determining whether a newborn will eventually develop leukemia




~1.4% of people will develop leukemia

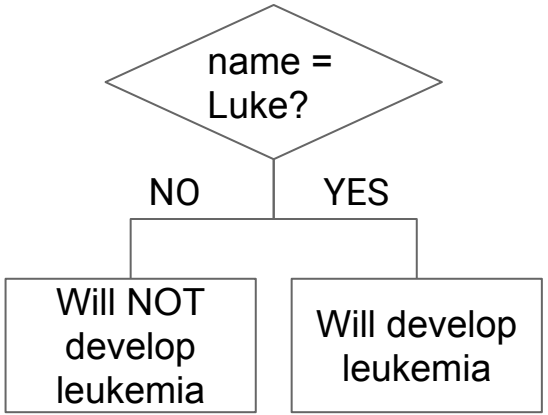
| | **leukemia** | **no leukemia** | **total** |
|---|---|---|---|
| "Luke" | 70 | 4,930 | 5,000 |
| not "Luke" | 13,930 | 981,070 | 99,500 |
| total | 14,000 | 986,000 | 1,000,000 |

5 out of every 1,000 babies named Luke

*Example from Data Science from Scratch, Ch 11*

# Pitfalls with using Accuracy

- Take the following decision tree for determining whether a newborn will eventually develop leukemia




| | leukemia | no leukemia | total |
|---|---|---|---|
| "Luke" | 70 | 4,930 | 5,000 |
| not "Luke" | 13,930 | 981,070 | 99,500 |
| total | 14,000 | 986,000 | 1,000,000 |

**Correctly predicts the result ~98% of the time! So this model must be good, right?**

# Alternatives to Accuracy

- Precision: The percentage of "positive" predictions that were correct
  - In the previous example, 70/5000 = 0.014
- Recall: The percentage of "positive" cases we correctly predict
  - In the previous example, 70/14000 = 0.005
- F1 Score: Harmonic mean of precision and recall
  - (2 * precision * recall) / (precision + recall)
  - In the previous example, 0.007

**All of these numbers are terrible for this model**

# k-Nearest Neighbors Demo

# Project Phase 2 - Modeling

- 5 different models
- 1 (or 2) must be from outside of class
- Must implement each model well –tune, train, evaluate, etc
- Must explain 3 things:
  - Why you chose the model for your particular problem
  - What work you did to use the model effectively
  - What did you learn from applying your model (both in terms of model effectiveness and what you learned about your data)
- Accuracy/performance is **not** important*, your explanations are

* this assumes you've convinced us you've done an effective job tuning/training your model