

# Tutorial 5 - Performance Metrics

Kevin Dick, PhD Candidate Biomedical Engineering  
Carleton University

Friday 9<sup>th</sup> October, 2020

# Disclaimer: Recorded Tutorials will be Publicly Posted

**Goal:** to create a companion series of applied machine learning tutorials for the 100MLB text, these tutorials will be publicly posted as a YouTube playlist.

## Privacy Preservation:

- Ask questions in the chat<sup>1</sup>
- Keep video off

**Note:** If the above *hinders your ability to learn  $\wedge$  violates your privacy*, please let me/Dr. Green know ASAP and video will be post-processed accordingly.

---

<sup>1</sup>I encourage unmuted/voice-based questions at any time, but know that this isn't explicitly privacy-preserving

# ML Weekly

## Recent news events from the ML community

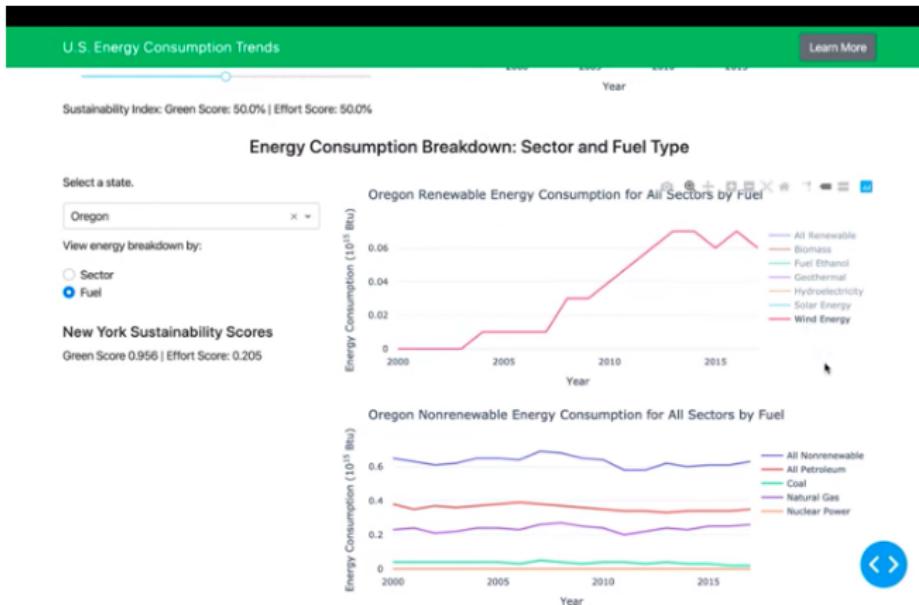
# ML Weekly

1. (ML) NeurIPS 2020 Accepted Papers
- 



# ML Weekly

1. (ML) NeurIPS 2020 Accepted Papers
2. (ML) Plotly Visualization



# ML Weekly

1. (ML) NeurIPS 2020 Accepted Papers
2. (ML) Plotly Visualization
3. (NLP) Reddit: thegentlemetre (GPT-3 Bot Spends a Week Replying on Reddit, Starts Talking About the Illuminati)



# Tutorial Intuition

Building an Intuition for the Concepts of this Tutorial

# How Good is my Model?

Welcome to the *metric zoo*!

There exists several dozens of model evaluation metrics and it is critical that one pairs the **right metric** for the **right problem**.

Universally, when learning about machine learning, you will

- **Classification-Type:** accuracy, precision, recall, specificity, F1-score, ROC, AUC, ...
- **Regression-Type:** MSE, MAE, RMSE, ...
- Rank-Aware Metrics
- Statistical-Type Metrics
- Computer Vision Metrics
- NLP Metrics
- Deep Learning-Related Metrics

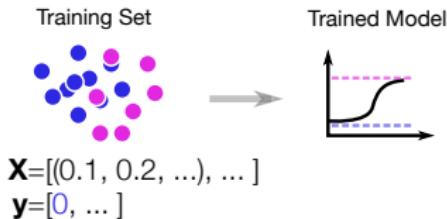
# How Good is my Model?

Training Set

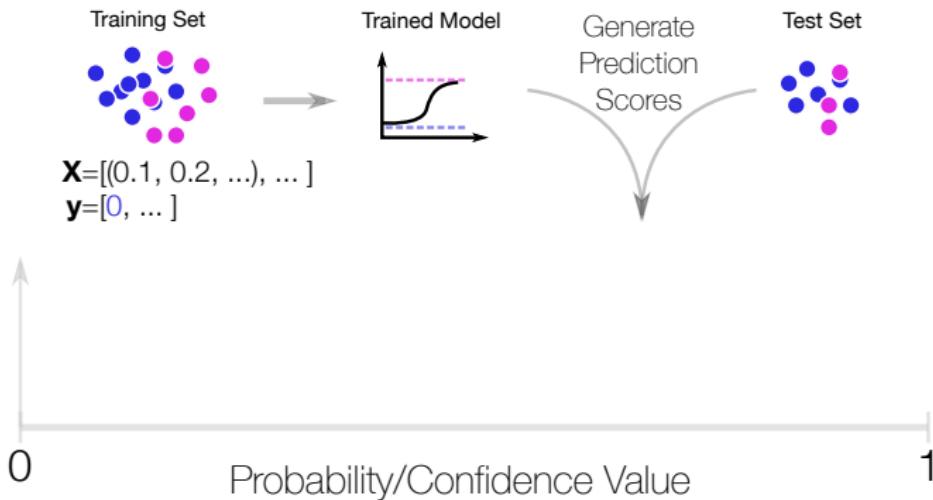


**X**=[(0.1, 0.2, ...), ... ]  
**y**=[0, ... ]

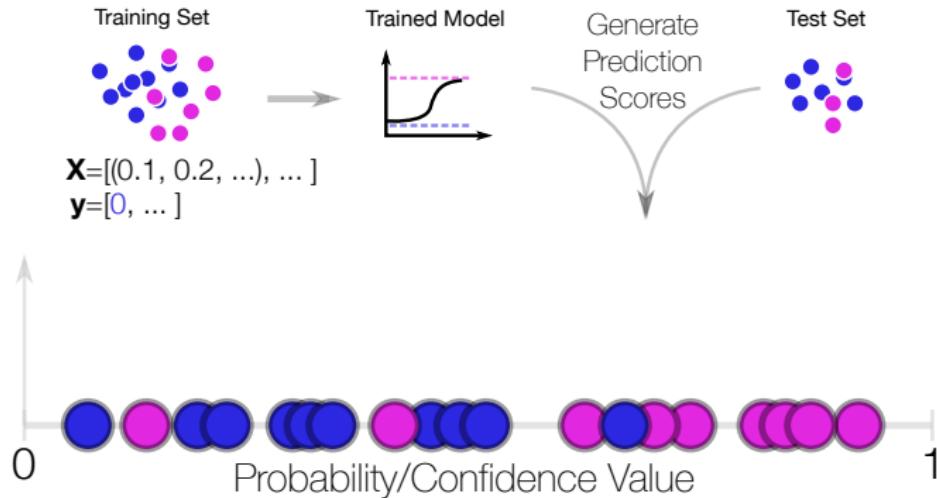
# How Good is my Model?



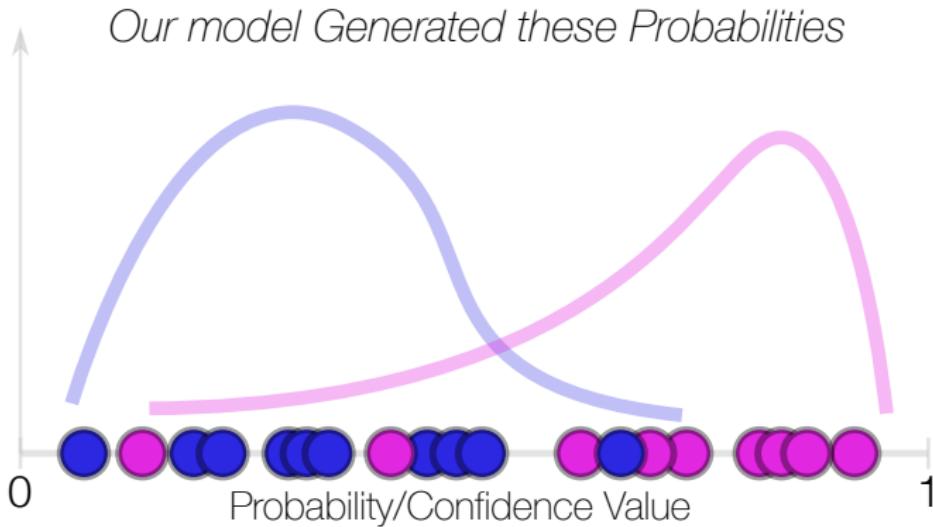
# How Good is my Model?



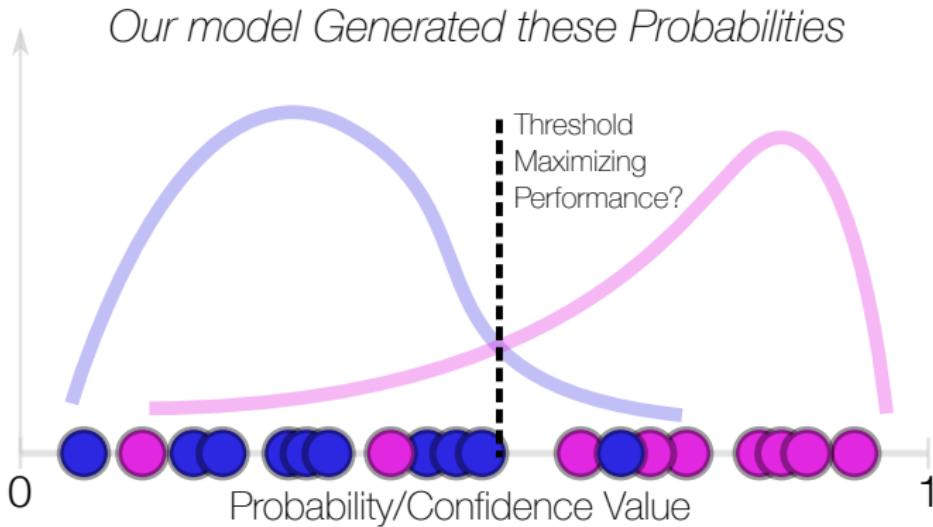
# How Good is my Model?



# How Good is my Model?

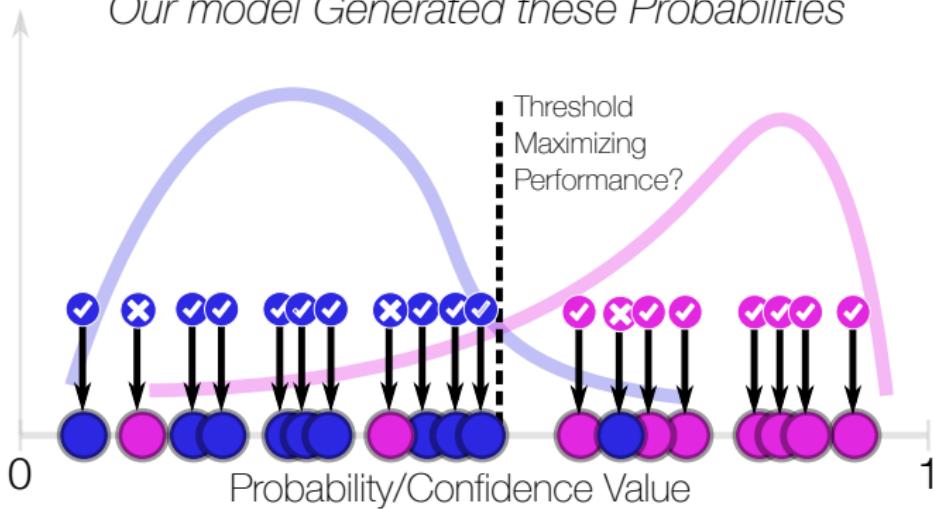


# How Good is my Model?

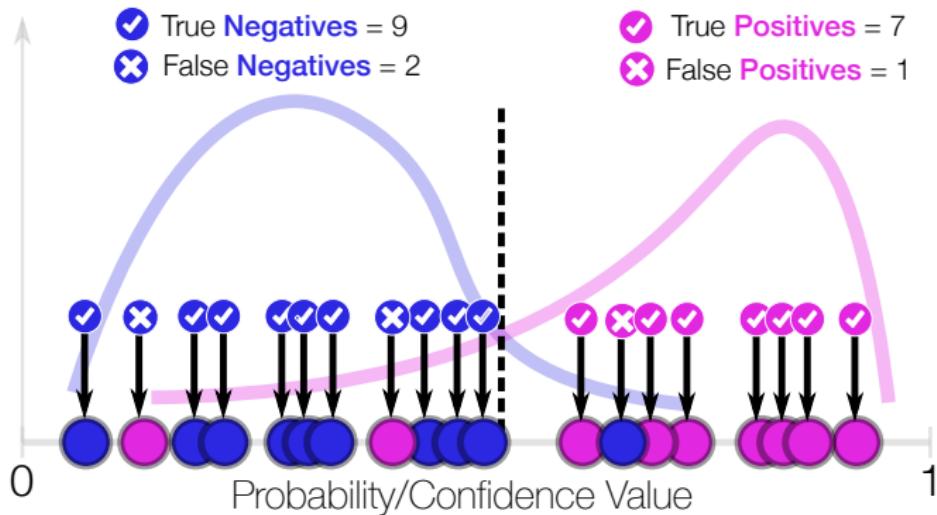


# How Good is my Model?

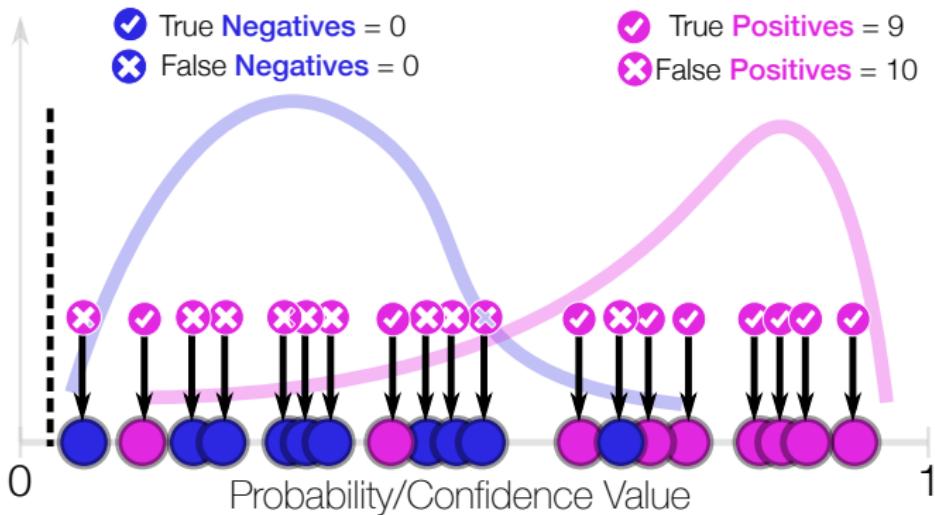
*Our model Generated these Probabilities*



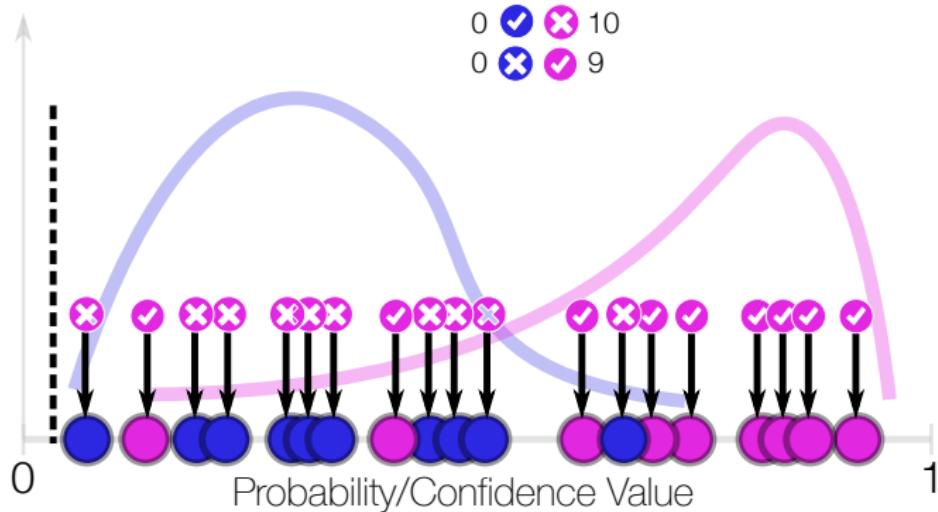
# How Good is my Model?



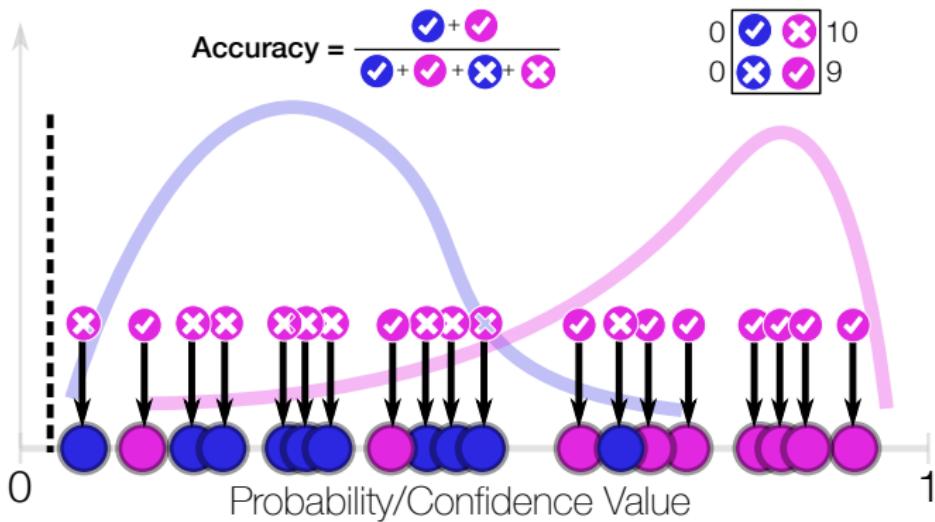
# How Good is my Model?



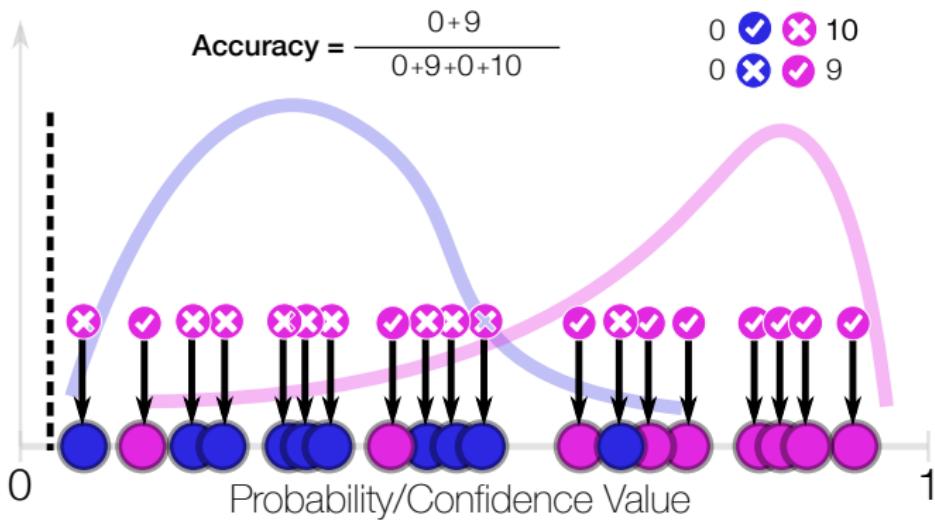
# How Good is my Model?



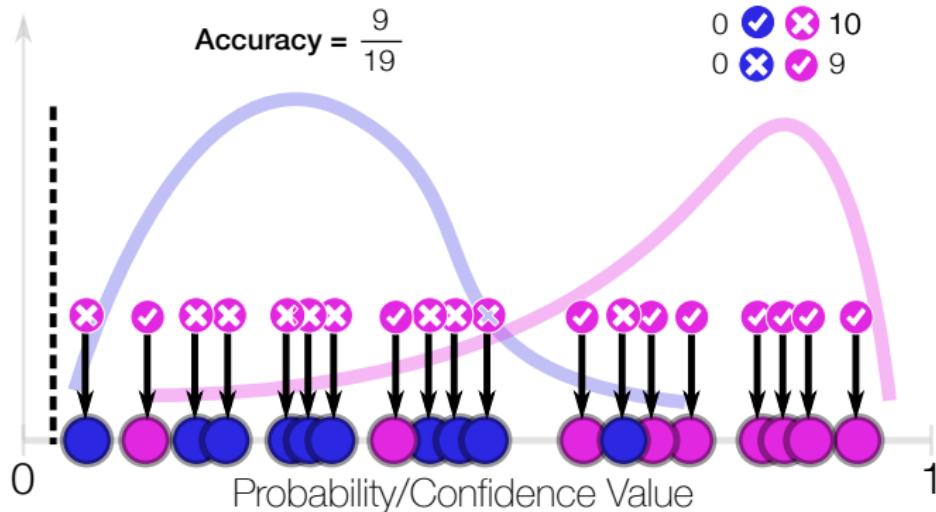
# How Good is my Model?



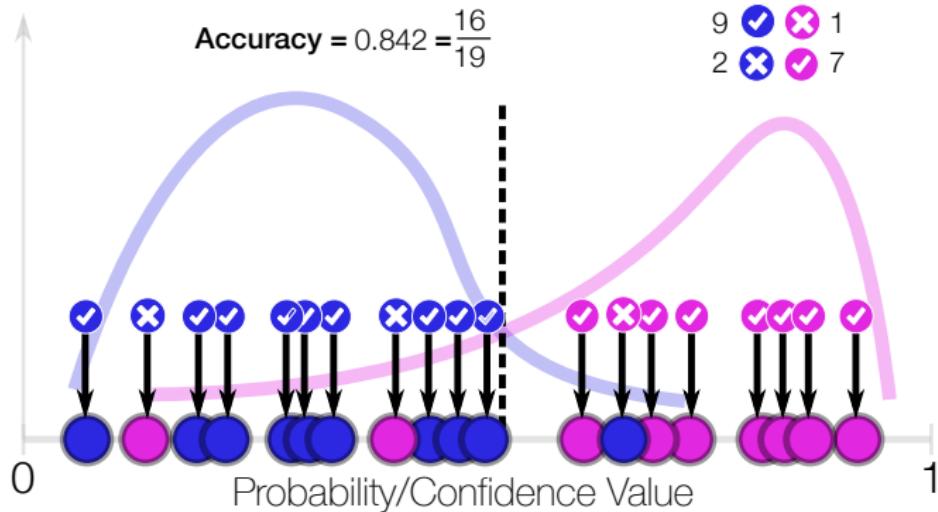
# How Good is my Model?



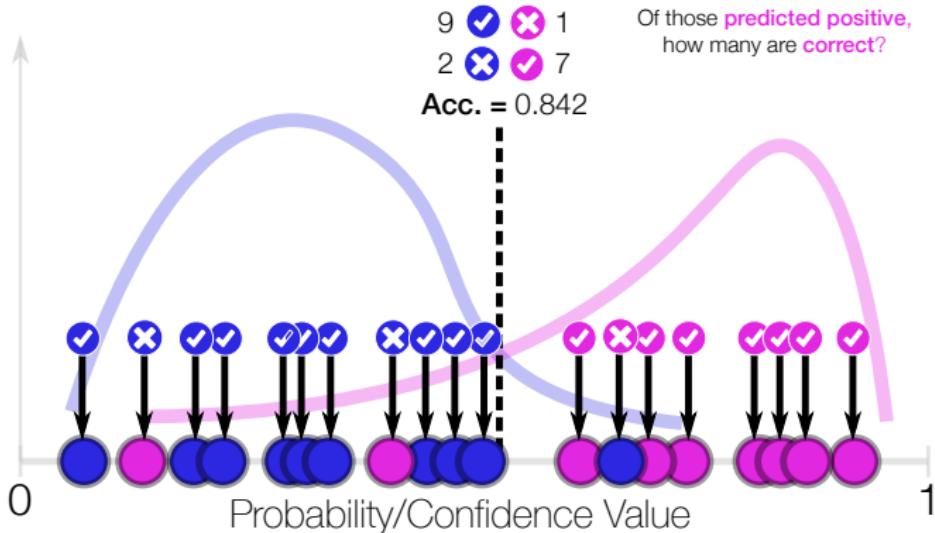
# How Good is my Model?



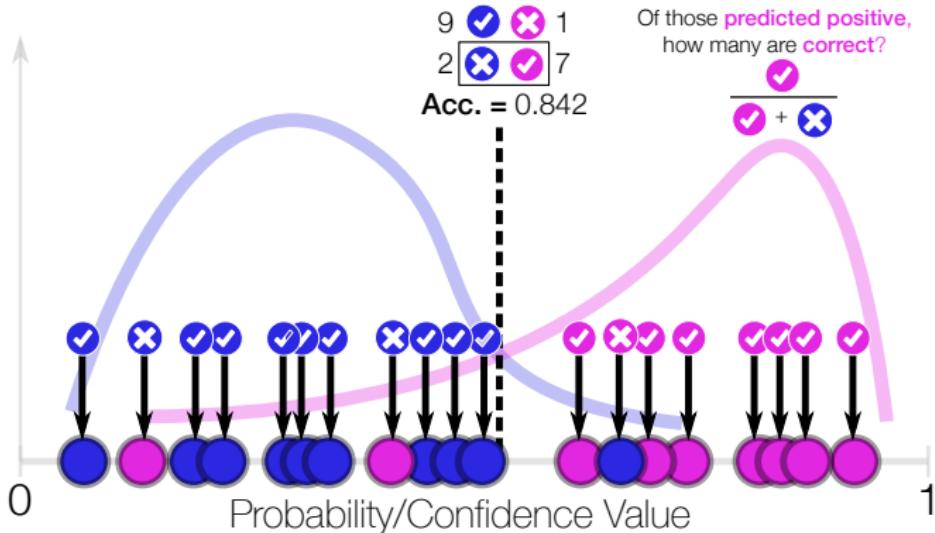
# How Good is my Model?



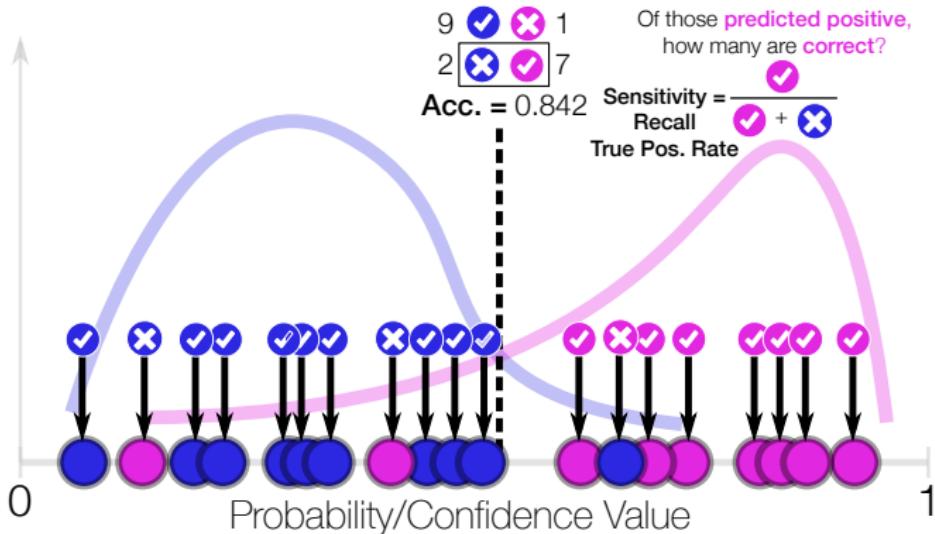
# How Good is my Model?



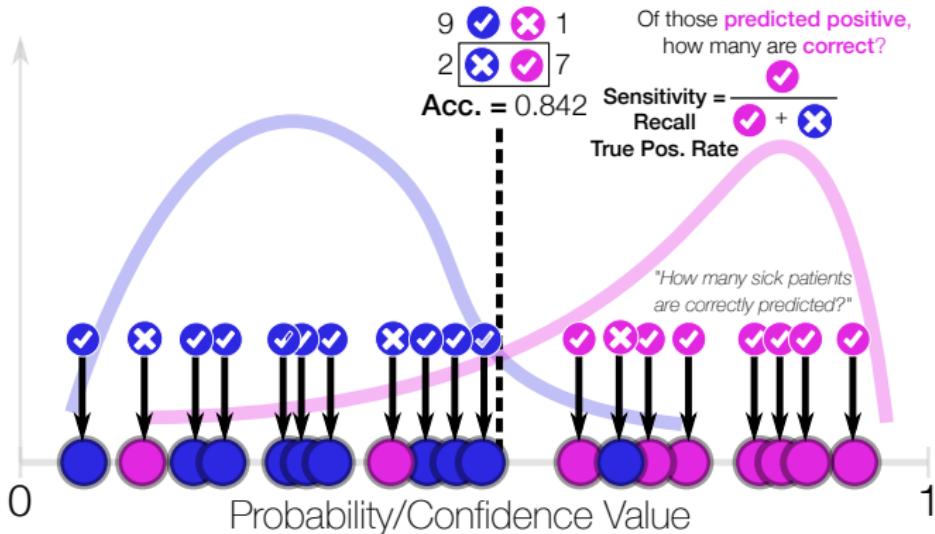
# How Good is my Model?



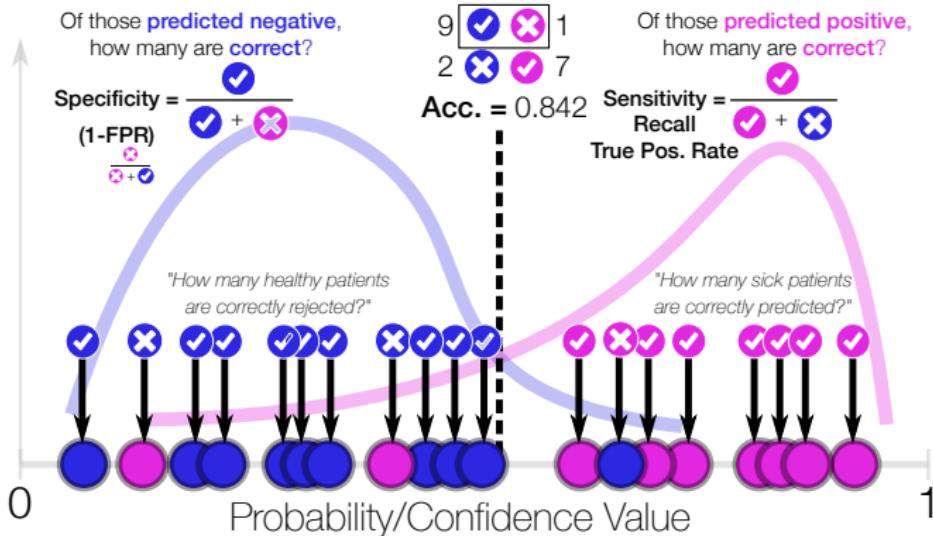
# How Good is my Model?



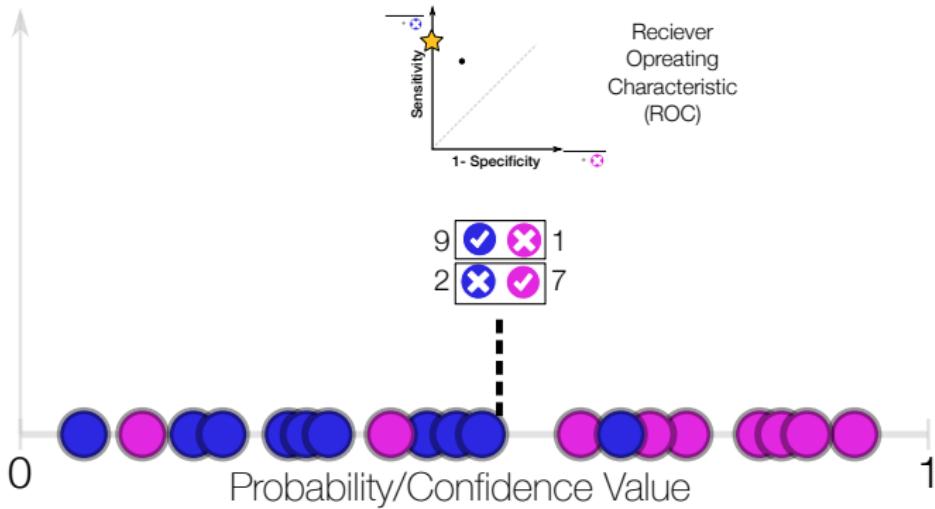
# How Good is my Model?



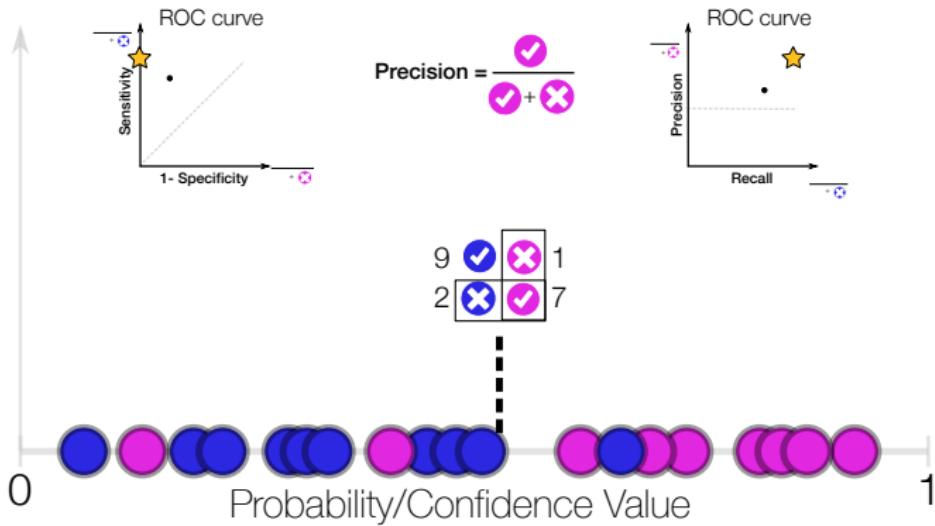
# How Good is my Model?



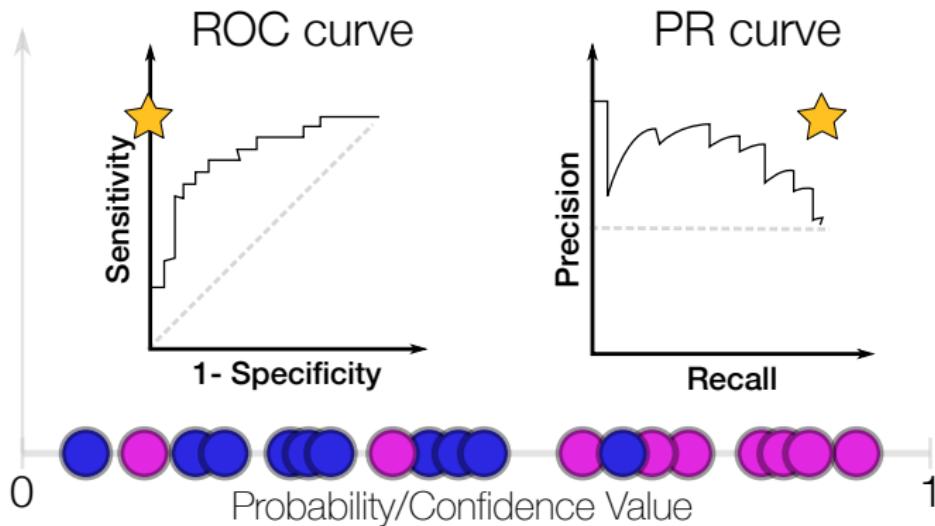
# How Good is my Model?



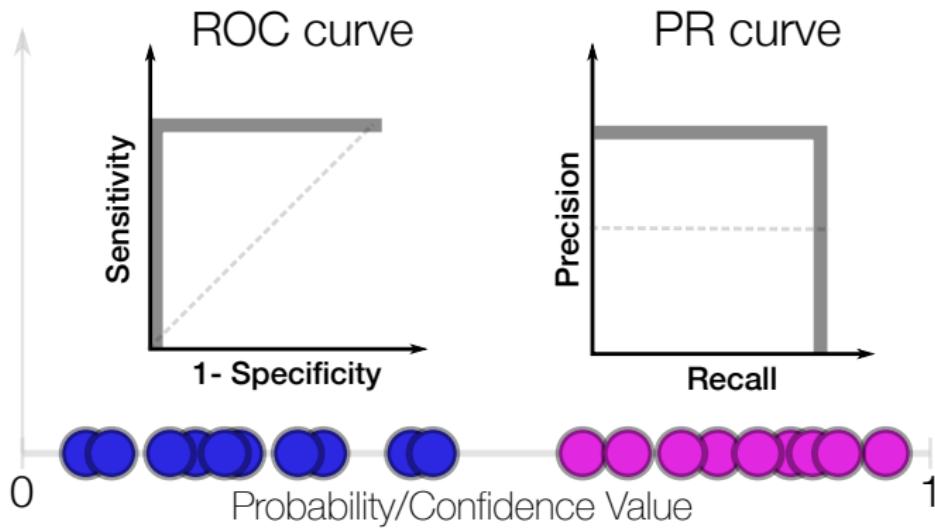
# How Good is my Model?



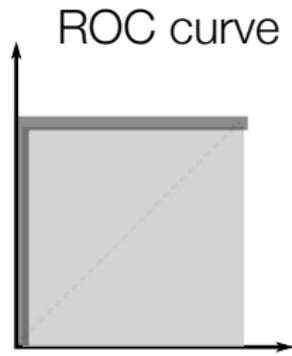
# How Good is my Model?



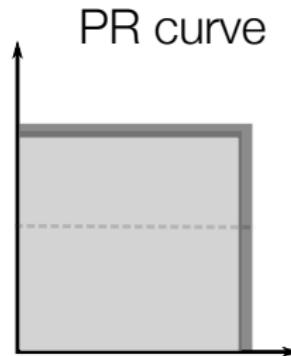
# How Good is my Model?



# How Good is my Model?



**AUC**  
**ROC = 1**



**AUC**  
**PR = 1**

# Animation: Variation in Threshold

# Animation: Improved Models by ROC Curve

# Animation: Improved Models by PR Curve

# Animation: Variation in Standard Deviation

# Animation: Variation in Class Imbalance

# Animation: Variation in Class Imbalance

# Into the Notebooks we Go...

We will cover one new notebook today!

## 1. Tutorial 5 - Performance Metrics

# Tutorial 5 - Performance Metrics

Kevin Dick, PhD Candidate Biomedical Engineering  
Carleton University

Friday 9<sup>th</sup> October, 2020