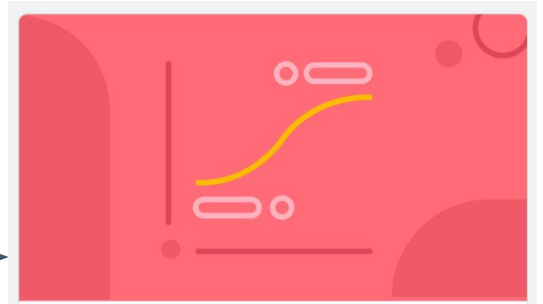
TECHCRUSH ARTIFICIAL INTELLIGENCE BOOTCAMP

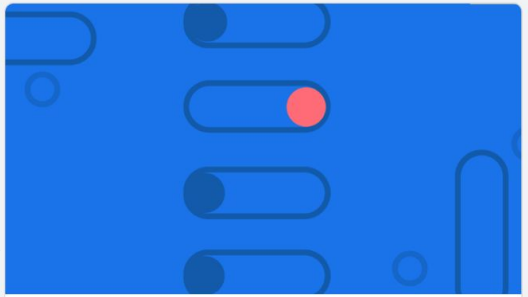
Facilitator: Hammed Obasekore
September 3st, 2025

Recap



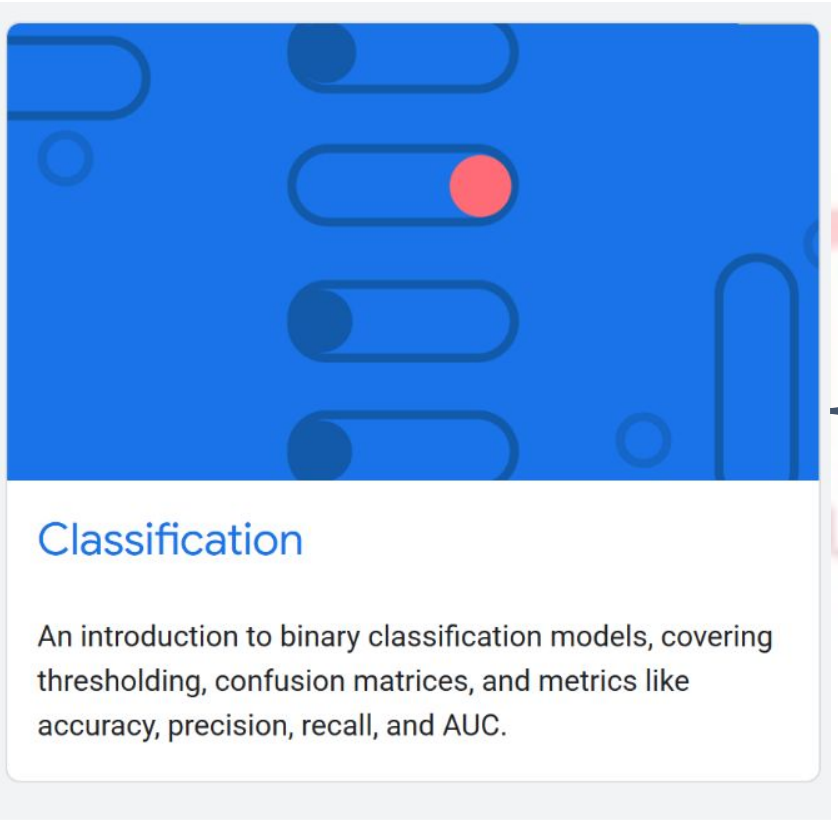
Logistic Regression

An introduction to logistic regression, where ML models are designed to predict the probability of a given outcome.



Classification

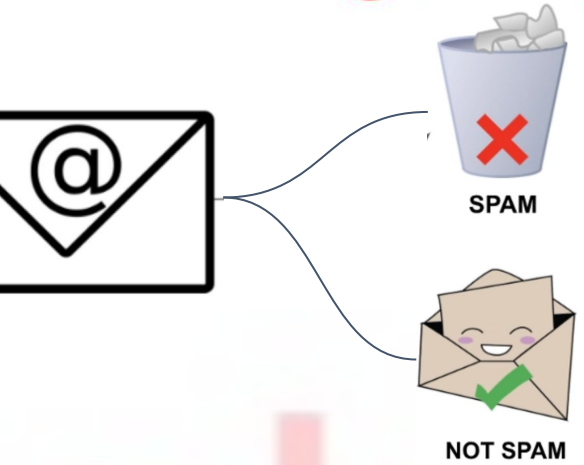
An introduction to binary classification models, covering thresholding, confusion matrices, and metrics like accuracy, precision, recall, and AUC.



Classification

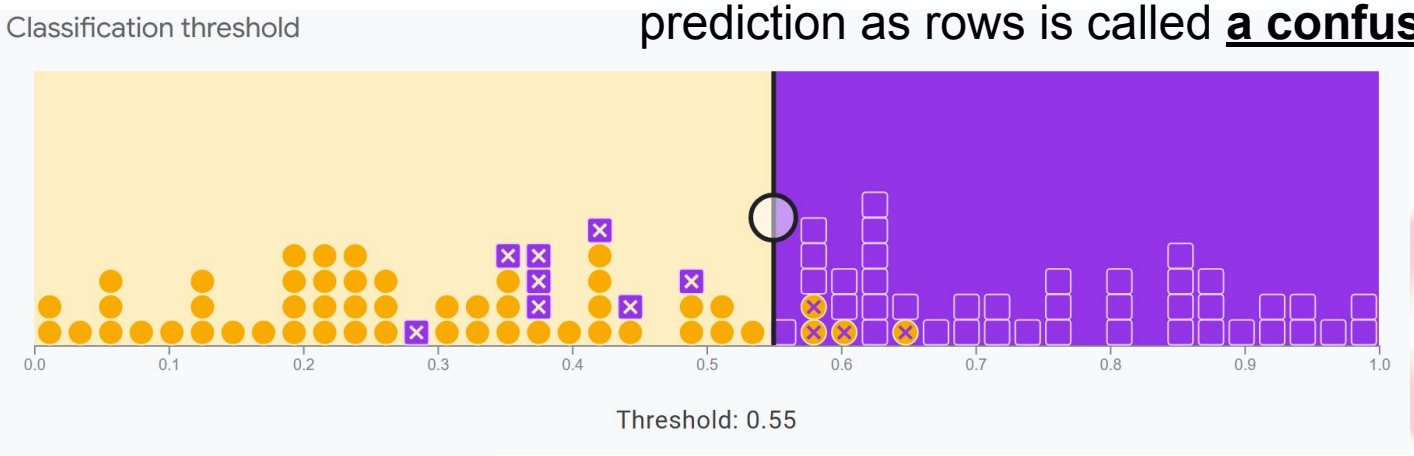
Binary Classification

Multi Class Classification



Classification: Thresholding & Confusion Matrix

A table with ground truth as columns and the model's prediction as rows is called **a confusion matrix**

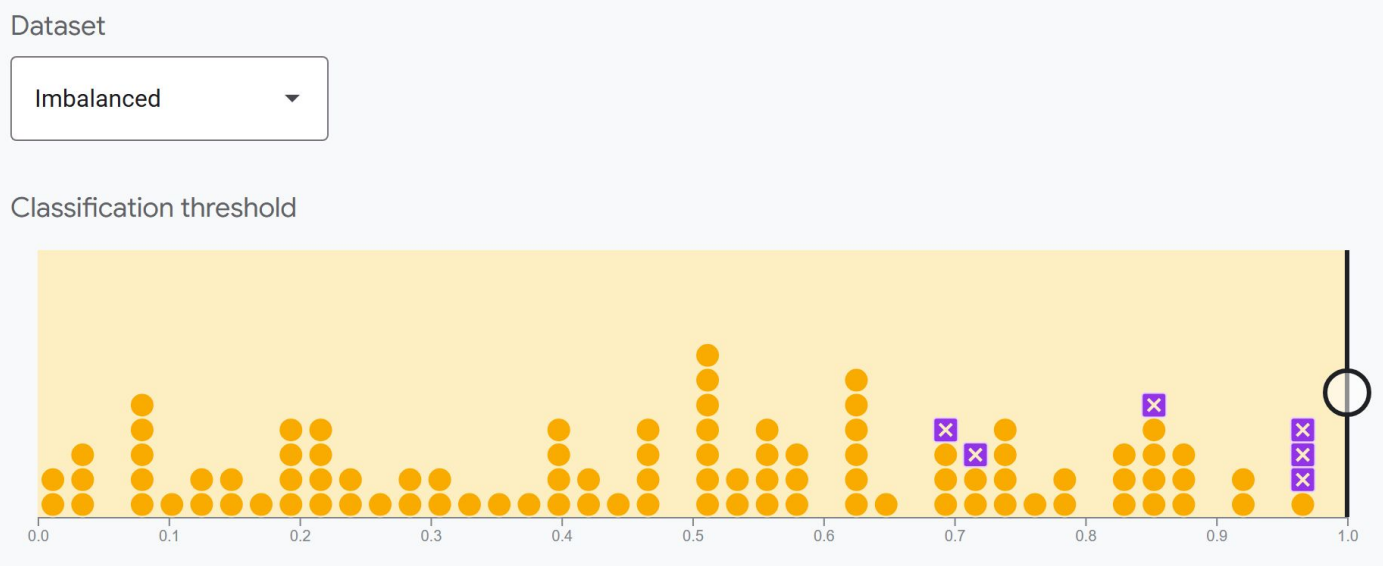


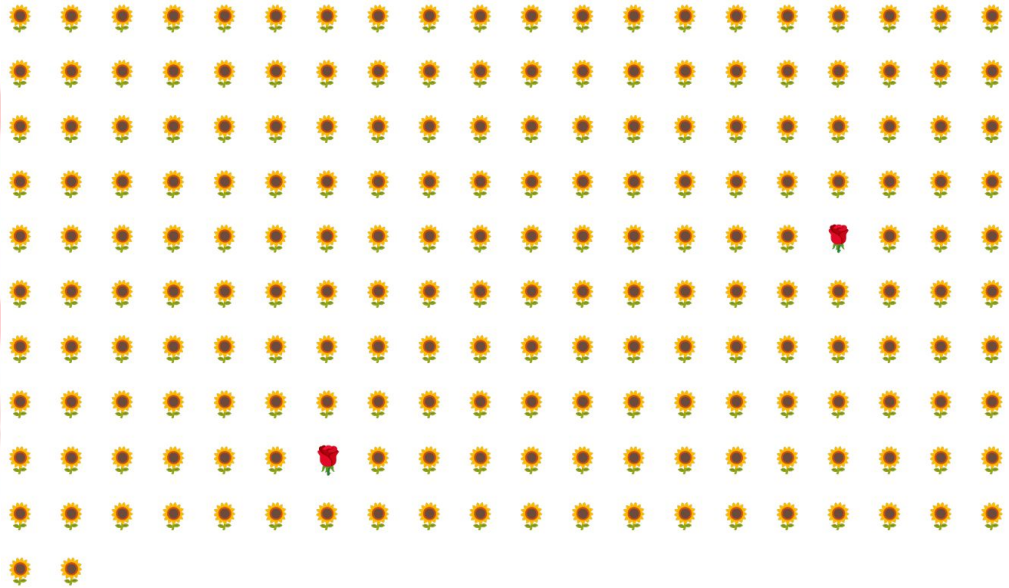
Confusion matrix

	Actually positive	Actually negative
Predicted positive	TP=40	FP=4
Predicted negative	FN=8	TN=47

	Actual positive	Actual negative
Predicted positive	**True positive (TP)**: A spam email correctly classified as a spam email. These are the spam messages automatically sent to the spam folder.	**False positive (FP)**: A not-spam email misclassified as spam. These are the legitimate emails that wind up in the spam folder.
Predicted negative	**False negative (FN)**: A spam email misclassified as not-spam. These are spam emails that aren't caught by the spam filter and make their way into the inbox.	**True negative (TN)**: A not-spam email correctly classified as not-spam. These are the legitimate emails that are sent directly to the inbox.

Classification: Data Imbalance





Classification: Accuracy, recall, precision, and related metrics

Accuracy is the proportion of **all classifications** that were **correct**, whether positive or negative.

$$\text{Accuracy} = \frac{\text{correct classifications}}{\text{total classifications}} = \frac{TP + TN}{TP + TN + FP + FN}$$

The true positive rate (TPR), or the proportion of all **actual positives** that were **classified correctly as positives**, is also known as recall.

$$\text{Recall (or TPR)} = \frac{\text{correctly classified actual positives}}{\text{all actual positives}} = \frac{TP}{TP + FN}$$

Precision is the proportion of all the model's positive classifications that are actually positive.

$$\text{Precision} = \frac{\text{correctly classified actual positives}}{\text{everything classified as positive}} = \frac{TP}{TP + FP}$$

Classification: Accuracy, recall, precision, and related metrics

The false positive rate (FPR) is the proportion of all actual negatives that were classified incorrectly as positives, also known as the probability of false alarm.

$$\text{FPR} = \frac{\text{incorrectly classified actual negatives}}{\text{all actual negatives}} = \frac{FP}{FP + TN}$$

The **F1 score** is the harmonic mean (a kind of average) of precision and recall.
Preferable to accuracy for **class-imbalanced datasets.**

$$\text{F1} = 2 * \frac{\text{precision} * \text{recall}}{\text{precision} + \text{recall}} = \frac{2\text{TP}}{2\text{TP} + \text{FP} + \text{FN}}$$

Classification: Accuracy, recall, precision, and related metrics

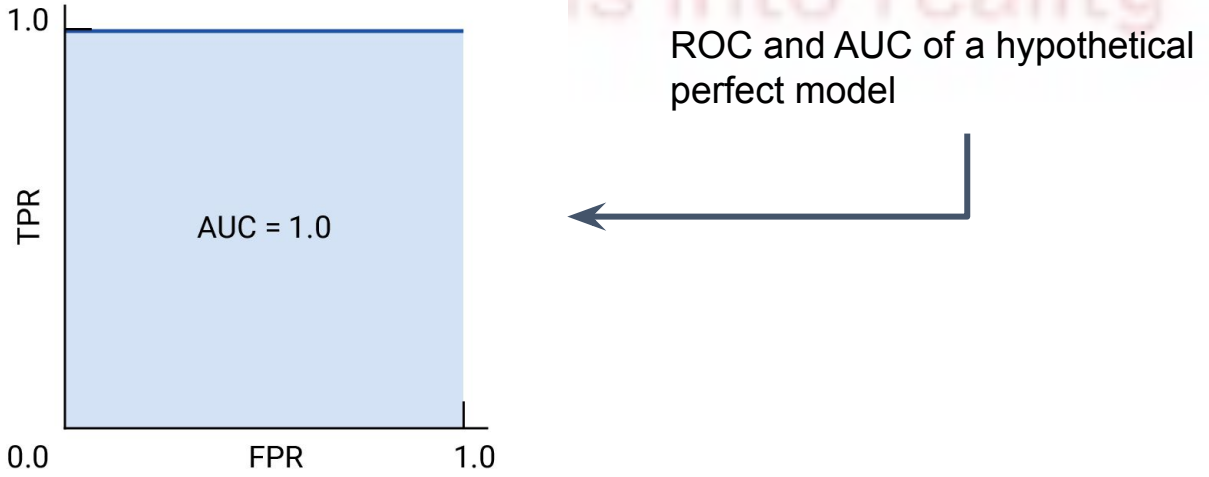
metric and tradeoffs

Metric	Guidance
Accuracy	Use as a rough indicator of model training progress/convergence for balanced datasets. For model performance, use only in combination with other metrics. Avoid for imbalanced datasets. Consider using another metric.
Recall (True positive rate)	Use when false negatives are more expensive than false positives.
False positive rate	Use when false positives are more expensive than false negatives.
Precision	Use when it's very important for positive predictions to be accurate.

Classification: : ROC and AUC

To evaluate a model's quality across all possible thresholds, you need different tools. Such as

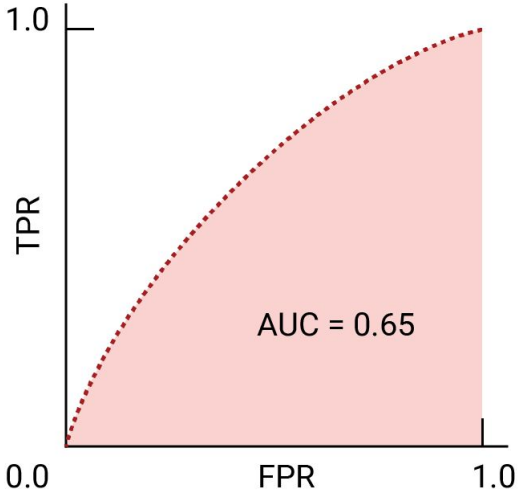
- Receiver-operating characteristic curve (ROC): a visual representation of model performance across all thresholds.
- Area under the curve (AUC): represents the probability that the model, if given a randomly chosen positive and negative example, will rank the positive higher than the negative.

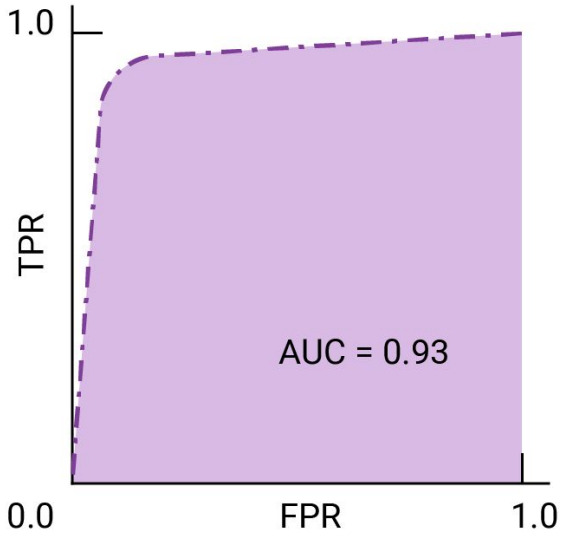


Classification: ROC and AUC

AUC and ROC for choosing model and threshold

AUC is a useful measure for comparing the performance of two different models, as long as the dataset is roughly balanced. The model with greater area under the curve is generally the better one.





The curve on the right, with a greater AUC, represents the better of the two models.

Questions?