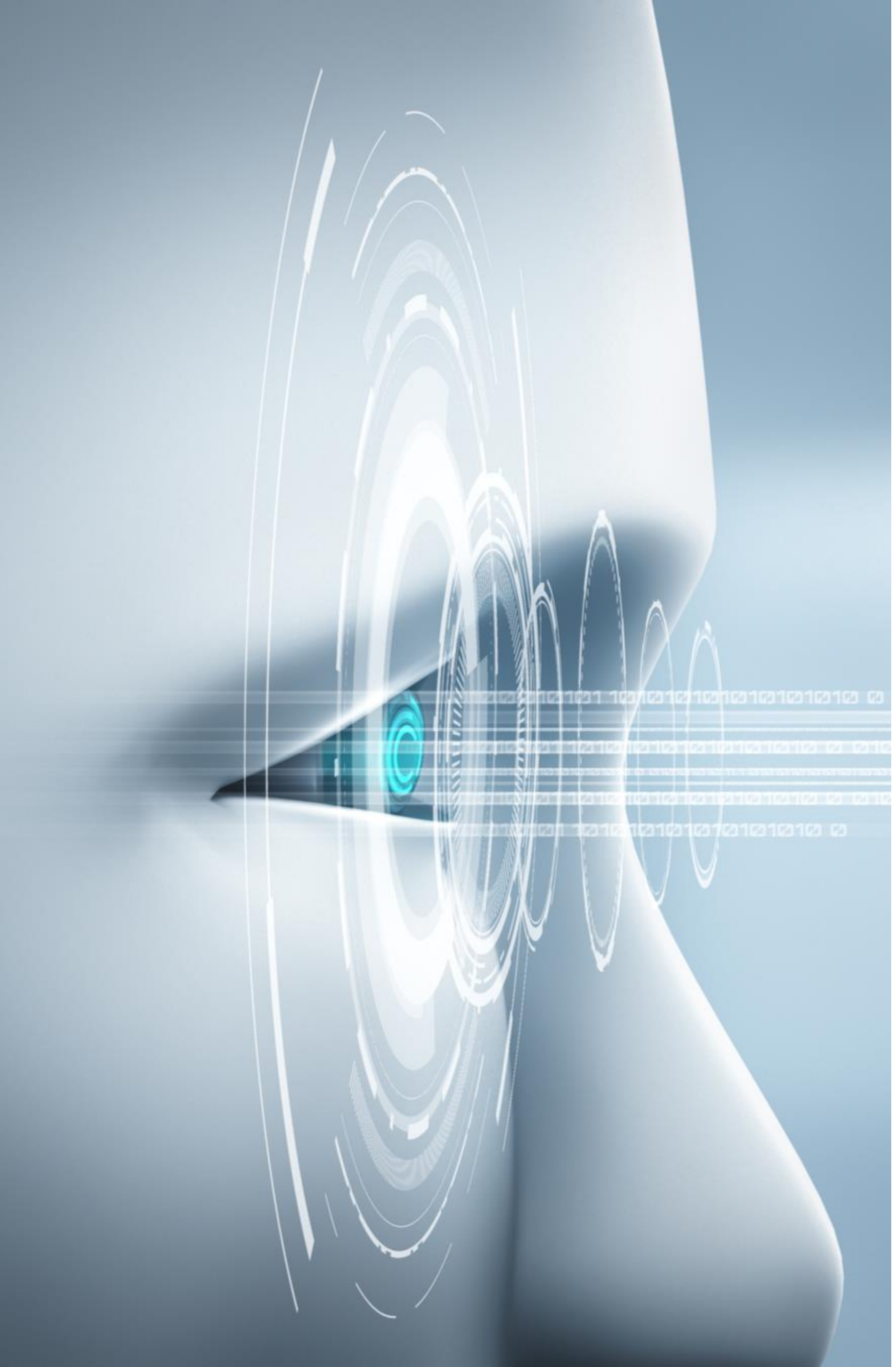


Computer Vision Applications

BY QUADEER SHAIKH

About me

Work Experience

- Risk Analyst
  - Morgan Stanley (Jan 2023 – Present)
- Data Science Intern
  - AkzoNobel Coatings International B.V. Netherlands (Feb 2022 – Dec 2022)
- Data Science Intern
  - EzeRx Health Tech Pvt. Ltd. (Jan 2022 – July 2022)
- Associate Engineer
  - Tata Communications Ltd. (July 2019 – Aug 2020)
- Network Automation and Analysis Engineer Intern
  - Cisco (June 2018 – July 2018)

Education

- M.Tech – Artificial Intelligence
  - NMIMS (2021 - 2023, currently pursuing)
- B.E. – Computer Engineering
  - Mumbai University (2015 - 2019)

Classification Models: Improvement and Usage in Video Processing Pipelines

Type I and Type II Error

Which is more dangerous ?

		Reality	
		True	**False**
Measured or Perceived	**True**	Correct	**Type 1 error** False Positive
	False	**Type 2 error** False Negative	Correct

Type I and Type II Errors: Context Matters

You decide to get tested for COVID-19 based on mild symptoms. There are two errors that could potentially occur:

- **Type I error (false positive):** the test result says you have coronavirus, but you actually don't.
- **Type II error (false negative):** the test result says you don't have coronavirus, but you actually do.

You build a model for predicting if a credit card user is going to default or not. The bank that uses this model might want to take some precautionary measures based on the prediction.

- **Type I error (false positive):** the model result says user will default, but he actually doesn't.
- **Type II error (false negative):** the model result says user will not default, but he actually does.

Confusion Matrix

Decide on the threshold value for classification based on the confusion matrix

		True Class	
		T	F
Acquired Class	Y	True Positives (TP)	False Positives (FP)
	N	False Negatives (FN)	True Negatives (TN)

$$\text{True Positive Rate (TPR)} = \frac{TP}{TP + FN}$$

$$\text{False Positive Rate (FPR)} = \frac{FP}{FP + TN}$$

$$\text{Accuracy (ACC)} = \frac{TP + TN}{TP + FP + TN + FN}$$

Different Metrics for Classification

Recall Sensitivity True positive rate (TPR)	$\frac{TP}{FN + TP} = \frac{TP}{P}$
False positive rate (FPR) False alarm rate	$\frac{FP}{TN + FP} = \frac{FP}{N}$
Specificity True negative rate (TNR)	$\frac{TN}{TN + FP} = \frac{TN}{N} = 1 - FPR$
Precision	$\frac{TP}{TP + FP}$
False negative rate (FNR)	$\frac{FN}{FN + TP} = \frac{FN}{P}$
Accuracy	$\frac{TP + TN}{P + N} = \frac{TP + TN}{TP + TN + FP + FN}$

ROC (Receiver Operating Characteristic) Curve (Binary Classification)

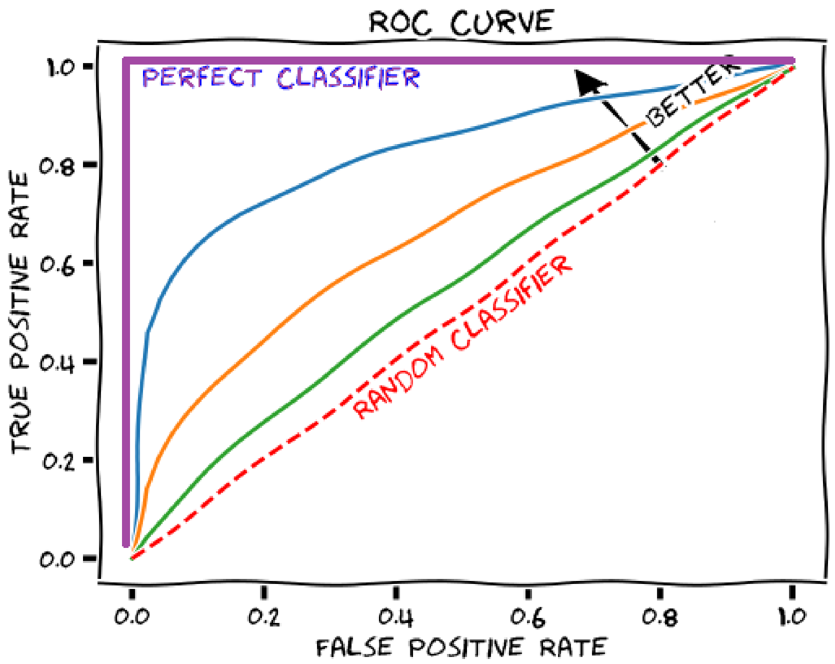
An **ROC curve** (**receiver operating characteristic curve**) is a graph showing the performance of a classification model at all classification thresholds. This curve plots two parameters:

- True Positive Rate
- False Positive Rate

predicted→ real↓	*Class_pos*	*Class_neg*
Class_pos	TP	FN
Class_neg	FP	TN

$$\text{TPR (sensitivity)} = \frac{TP}{TP + FN}$$

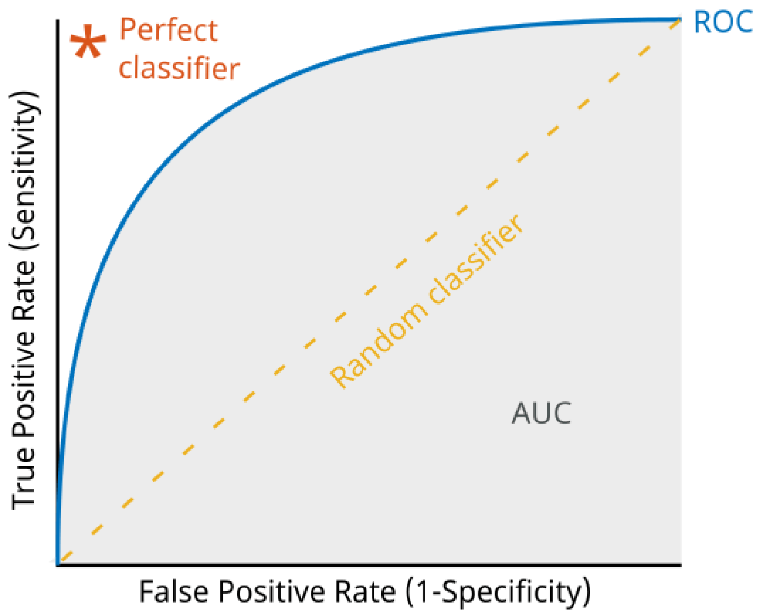
$$\text{FPR (1-specificity)} = \frac{FP}{TN + FP}$$



AUC (Area Under ROC Curve)

AUC provides an aggregate measure of performance across all possible classification thresholds.

AUC ranges in value from 0 to 1. A model whose predictions are 100% wrong has an AUC of 0.0; one whose predictions are 100% correct has an AUC of 1.0.



Losses in Classification Models

Binary Cross Entropy

$$L_{BCE} = -\frac{1}{n}\sum_{i=1}^{n}(Y_i \cdot \log \hat{Y}_i + (1 - Y_i) \cdot \log(1 - \hat{Y}_i))$$

Categorical Cross Entropy

$$\text{Loss} = -\sum_{i=1}^{\substack{\text{output}\\\text{size}}} y_i \cdot \log \hat{y}_i$$

Why should one monitor validation loss while saving the best model instead of accuracy ?

Which Model is better ?

Model 1		P(A)	P(B)	P(C)		GT(A)	GT(B)	GT(C)		Actual label	Predicted Label
	1	0.55	0.35	0.1		1	0	0		A	A
	2	0.3	0.5	0.2		0	1	0		B	B
	3	0.6	0.35	0.05		1	0	0		A	A
	4	0.3	0.3	0.4		0	0	1		C	C

Model 2		P(A)	P(B)	P(C)		GT(A)	GT(B)	GT(C)		Actual label	Predicted Label
	1	0.8	0.1	0.1		1	0	0		A	A
	2	0.5	0.4	0.1		0	1	0		B	A
	3	0.75	0.2	0.05		1	0	0		A	A
	4	0.05	0.2	0.75		0	0	1		C	C

Which Model is better ?

Space for Quadeer’s lethal mathematical skillzzzz

How to get more confident predictions ?

Different Ensembling techniques

1. Bagging
2. Boosting
3. Stacking
4. Voting
   i. Hard Voting: Takes the mode of predictions of different classifiers
   ii. Soft Voting: Takes the probability average of different classifiers and then decides a class

Video Processing Pipelines using Classification Models

Naïve Video Processing Pipeline

1. Give a video input source
2. Read each frame of the video
3. Classify each frame in the video (Bottleneck)
4. Display the prediction on the video

Video Processing Pipelines using Classification Models

MULTI THREADING (AKA CONCURRENT PROCESSING)

1. Multiple segments of a program/process is created to speed up the task at hand
2. Should be used when your task is I/O bound i.e. receiving an input continuously, processing it and then displaying it is a bottleneck.
3. E.g. You type sentences in a word file, and your spell check runs concurrently without blocking you from typing the next word.
4. E.g. Your model classifies each frame of the video and your program should not have to wait for fetching the next video frame thus blocking and slowing down the whole process

MULTIPROCESSING

1. A program/process is subdivided into multiple processes that can speed up a task by running this multiple subdivided processes on different processors/CPU cores.
2. Should be used when your task is computationally heavy and requires a lot of processors for computation.
3. E.g. Utilizing multiple cores of a GPU to train your deep learning models

Thank you

For any queries drop an email at: quadeershaikh15.8@gmail.com