

# FROM ACCURACY TO $R^2$ : THE ESSENTIAL MACHINE LEARNING METRICS YOU MUST KNOW



Prepared By :  
AKASH C

# WHAT ARE MACHINE LEARNING METRICS?

- IN MACHINE LEARNING, A METRIC IS A NUMERICAL SCORE THAT MEASURES HOW WELL A MODEL'S PREDICTIONS MATCH THE ACTUAL RESULTS.
- JUST LIKE MARKS IN SCHOOL OR SCORES IN CRICKET, METRICS TELL US IF THE MODEL IS PERFORMING WELL OR NEEDS IMPROVEMENT.
- A METRIC CONVERTS THE DIFFERENCE BETWEEN PREDICTED OUTPUTS AND ACTUAL VALUES INTO A SINGLE MEASURABLE VALUE, MAKING IT EASIER TO EVALUATE, COMPARE, AND IMPROVE MODELS.



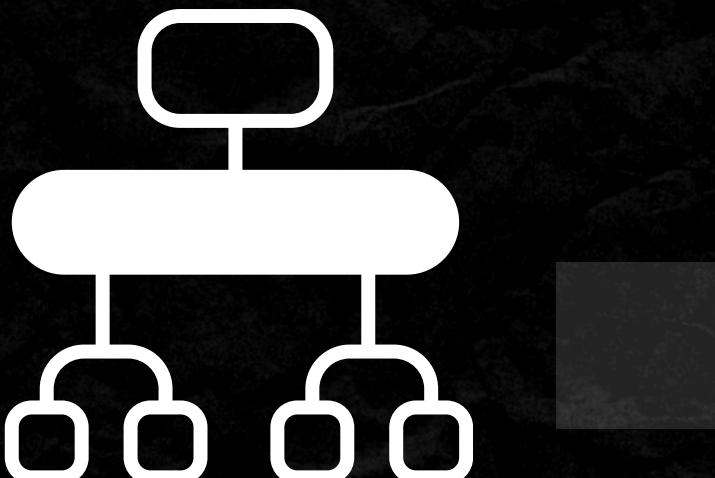
# WHY DO WE NEED METRICS?

- MODEL PERFORMANCE EVALUATION - MEASURE HOW WELL A MODEL HANDLES UNSEEN DATA AND DETECT UNDER/OVERFITTING.
- ERROR ANALYSIS & OPTIMIZATION - PINPOINT ERRORS, REFINE FEATURES, TUNE PARAMETERS, AND MATCH DOMAIN NEEDS.
- MODEL COMPARISON & DECISION-MAKING - RANK MODELS, ADJUST THRESHOLDS, AND SELECT THE ONE THAT MEETS BUSINESS GOALS.



# CLASSIFICATION METRICS

- USED WHEN TARGET VARIABLE = CATEGORIES (YES/NO, SPAM/NOT SPAM).
- COMMON METRICS:
  - **ACCURACY** → % OF CORRECT PREDICTIONS. (BEST FOR BALANCED DATA)
  - **PRECISION** → % OF PREDICTED POSITIVES, HOW MANY ARE CORRECT. (USEFUL WHEN FALSE POSITIVES ARE COSTLY, E.G., SPAM MAILS)
  - **RECALL (SENSITIVITY)** → % OF ACTUAL POSITIVES, HOW MANY MODEL DETECTED. (USEFUL WHEN MISSING POSITIVES IS COSTLY, E.G., DISEASE DIAGNOSIS)
  - **F1-SCORE** → BALANCE BETWEEN PRECISION & RECALL.
  - **ROC-AUC** → MODEL'S ABILITY TO SEPARATE CLASSES AT DIFFERENT THRESHOLDS.



# REGRESSION METRICS

- USED WHEN TARGET VARIABLE = CONTINUOUS VALUES (PRICE, MARKS, TEMPERATURE).
- COMMON METRICS:
  - MAE (MEAN ABSOLUTE ERROR) → AVERAGE SIZE OF ERRORS, TREATS ALL EQUALLY.
  - MSE (MEAN SQUARED ERROR) → PUNISHES LARGE ERRORS MORE.
  - RMSE (ROOT MEAN SQUARED ERROR) → SAME AS MSE BUT IN ACTUAL UNIT SCALE.
  - $R^2$  SCORE (COEFFICIENT OF DETERMINATION) → HOW MUCH VARIANCE IN DATA MODEL CAN EXPLAIN.



# CHOOSING THE RIGHT METRIC

- CLASSIFICATION:

- BALANCED DATASET → ACCURACY
- MINIMIZE FALSE POSITIVES → PRECISION
- MINIMIZE FALSE NEGATIVES → RECALL
- NEED BALANCE → F1-SCORE
- FOR THRESHOLD COMPARISON → ROC-AUC

- REGRESSION:

- EQUAL ERROR WEIGHT → MAE
- LARGE ERRORS CRITICAL → MSE / RMSE
- EXPLAIN VARIATION → R<sup>2</sup> SCORE



# CONCLUSION

- ML METRICS ACT AS THE SCOREBOARD TO EVALUATE HOW WELL MODELS PERFORM.
- THE RIGHT METRIC CHOICE DEPENDS ON THE PROBLEM TYPE, DATASET, AND DOMAIN GOALS.
- OFTEN, USING A COMBINATION OF METRICS GIVES A MORE COMPLETE PERFORMANCE PICTURE.
- METRICS HELP IN FAIRLY COMPARING MODELS AND MAKING BETTER TRADE-OFF DECISIONS.
- ULTIMATELY, THEY ENSURE MODELS ARE NOT JUST ACCURATE BUT ALSO USEFUL AND RELIABLE IN REAL-WORLD APPLICATIONS.

