SNS COLLEGE OF ENGINEERING
(Autonomous)
DEPARTMENT OF ELECTRONICS AND COMMUNICATION ENGINEERING

Artificial Intelligence & Machine Learning

Testing Machine Learning Algorithms

Prepared by,
P.Ramya
Assistant Professor/ECE
SNS College of Engineering

What is the goal of ML testing?

- Quality Assurance

- Software testing

The purpose of machine learning testing is, first of all, to ensure that this learned logic will remain consistent, no matter how many times we call the program.

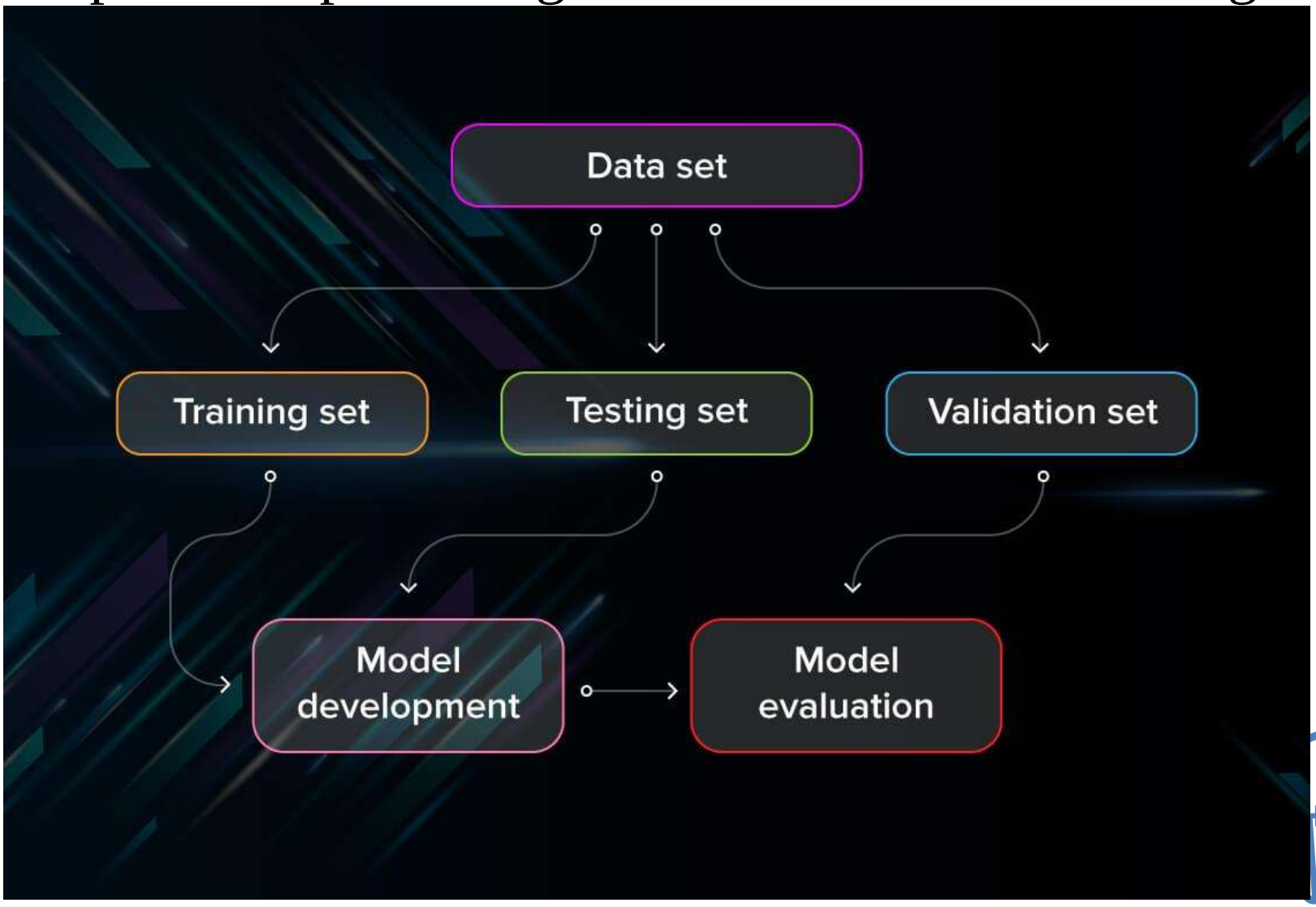
Model evaluation in machine learning testing

- **Unit tests.** The program is broken down into blocks, and each element (unit) is tested separately.
- **Regression tests**. They cover already tested software to see if it doesn't suddenly break.
- **Integration tests.** This type of testing observes how multiple components of the program work together.

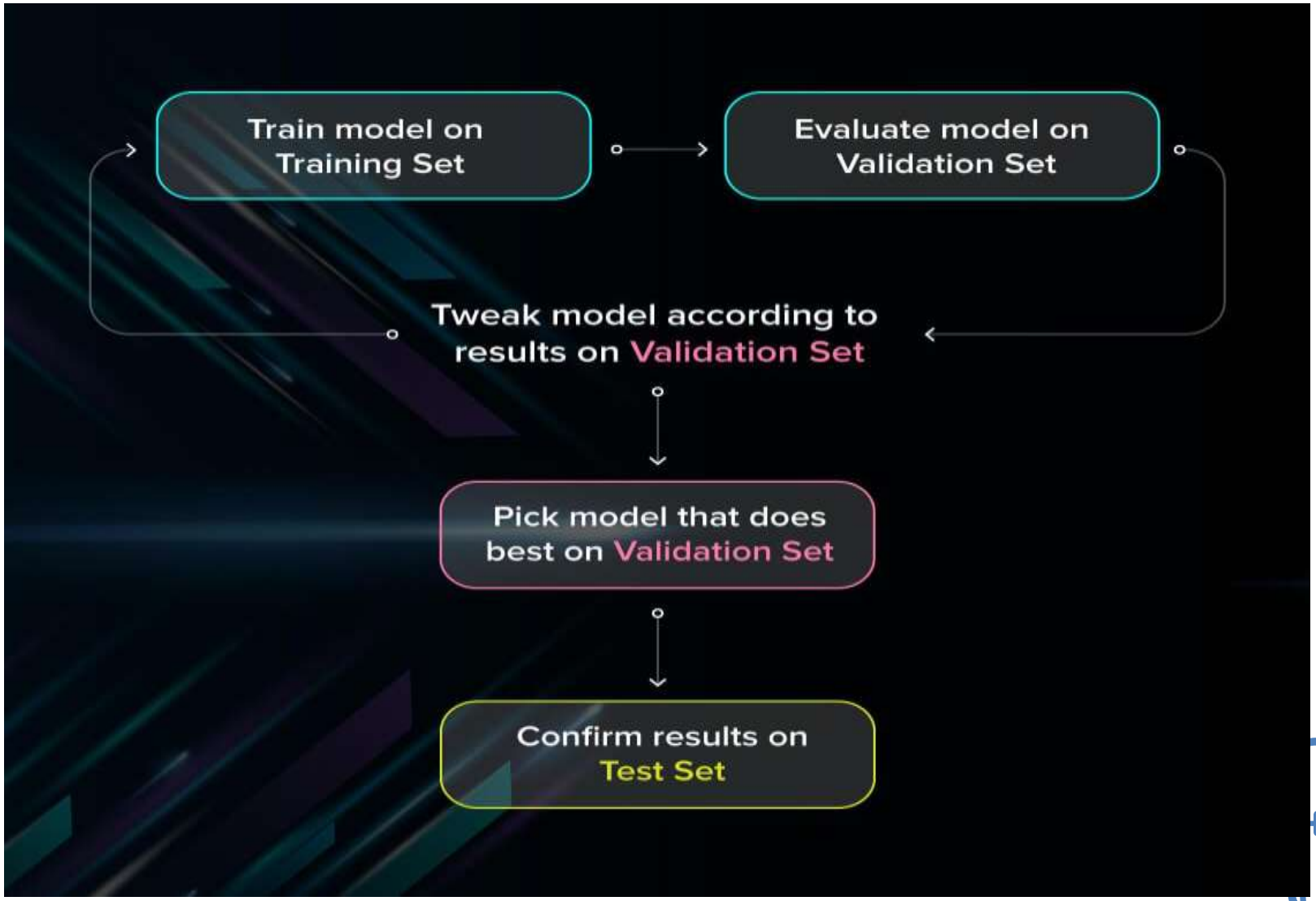
Contd…

Every ML model needs not only to be tested but evaluated. Your model should generalize well. This is not what we usually understand by testing, but evaluation is needed to make sure that the performance is satisfactory.

A Simple Example of High Dimensional Data Cursing Us



Contd…


Train model on Training Set
Evaluate model on Validation Set
Tweak model according to results on Validation Set
Pick model that does best on Validation Set
Confirm results on Test Set

Cross-validation

Cross-validation is a model evaluation technique that can be performed even on a limited dataset. The training set is divided into small subsets, and the model is trained and validated on each of these samples.

- k-fold cross-validation
- Leave-one-out cross-validation

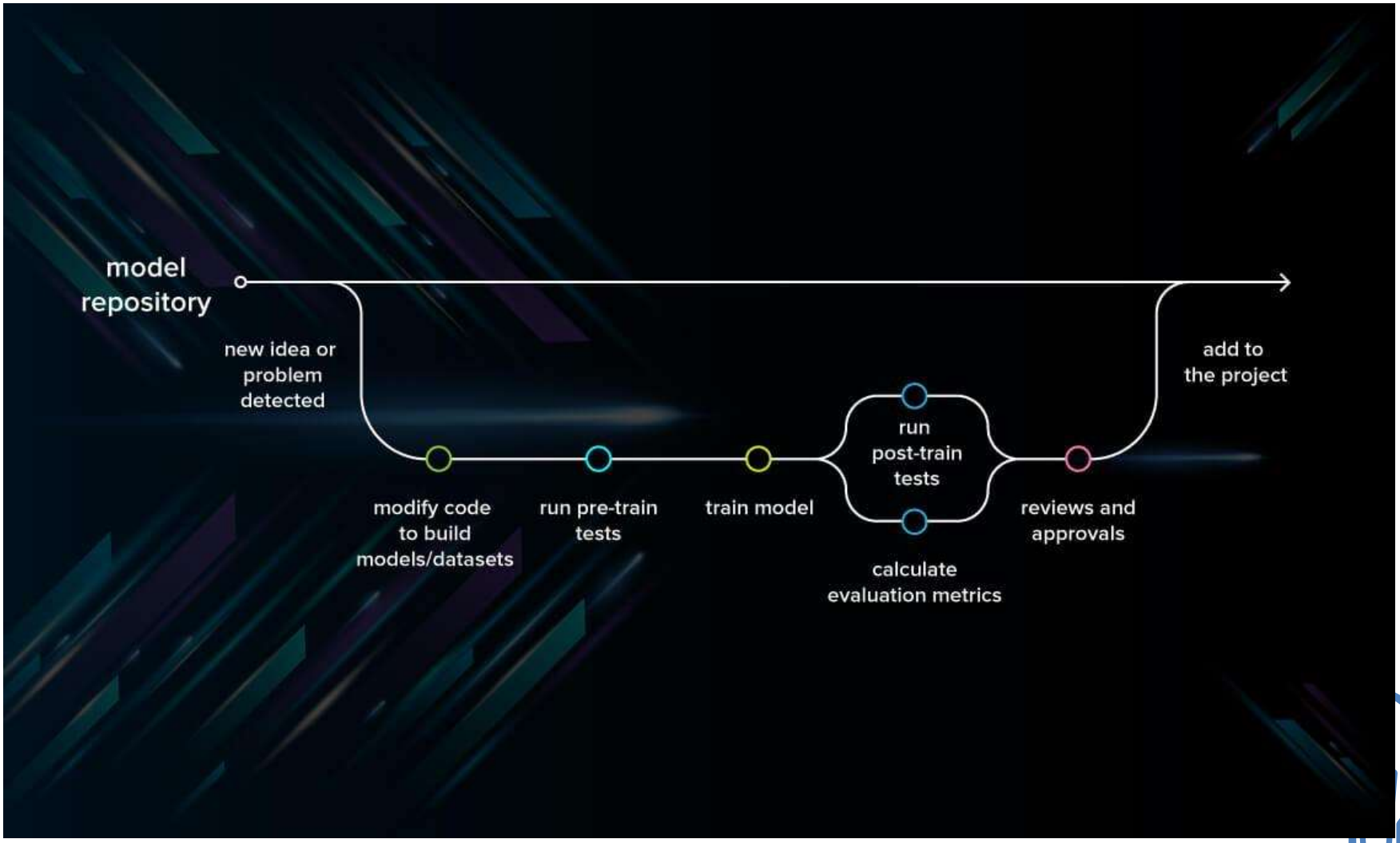
Evaluate models using metrics

- Accuracy
- Loss
- Precision
- Recall
- Confusion Matrix

How to write model tests?

- Pre-train tests
- Post-train tests
  - Invariance tests
  - Directional expectation tests
  - Minimum functionality tests

Model development pipeline



Conclusion

Performing ML tests is necessary if you care about the quality of the model. ML testing has a couple of peculiarities: it demands that you test the quality of data, not just the model, and go through a couple of iterations adjusting the hyperparameters to get the best results. However, if you perform all the necessary procedures, you can be sure of its performance.


thank
you