



Testing Machine Learning Algorithms

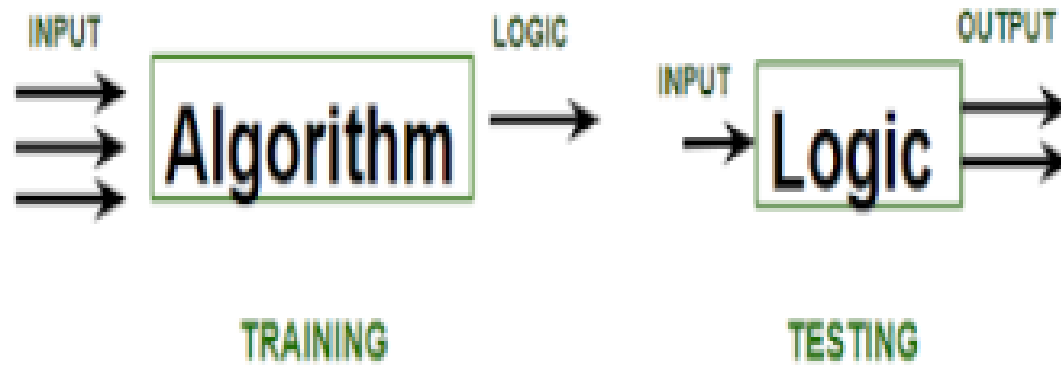
- In machine learning, model testing is referred to as the process where the **performance of a fully trained model** is evaluated on a testing set.
- Model evaluation in machine learning testing
 - Unit tests:
 - ✓ The program is broken down into blocks, and each element (unit) is tested separately.
 - ✓ To Check the correctness of individual model components.

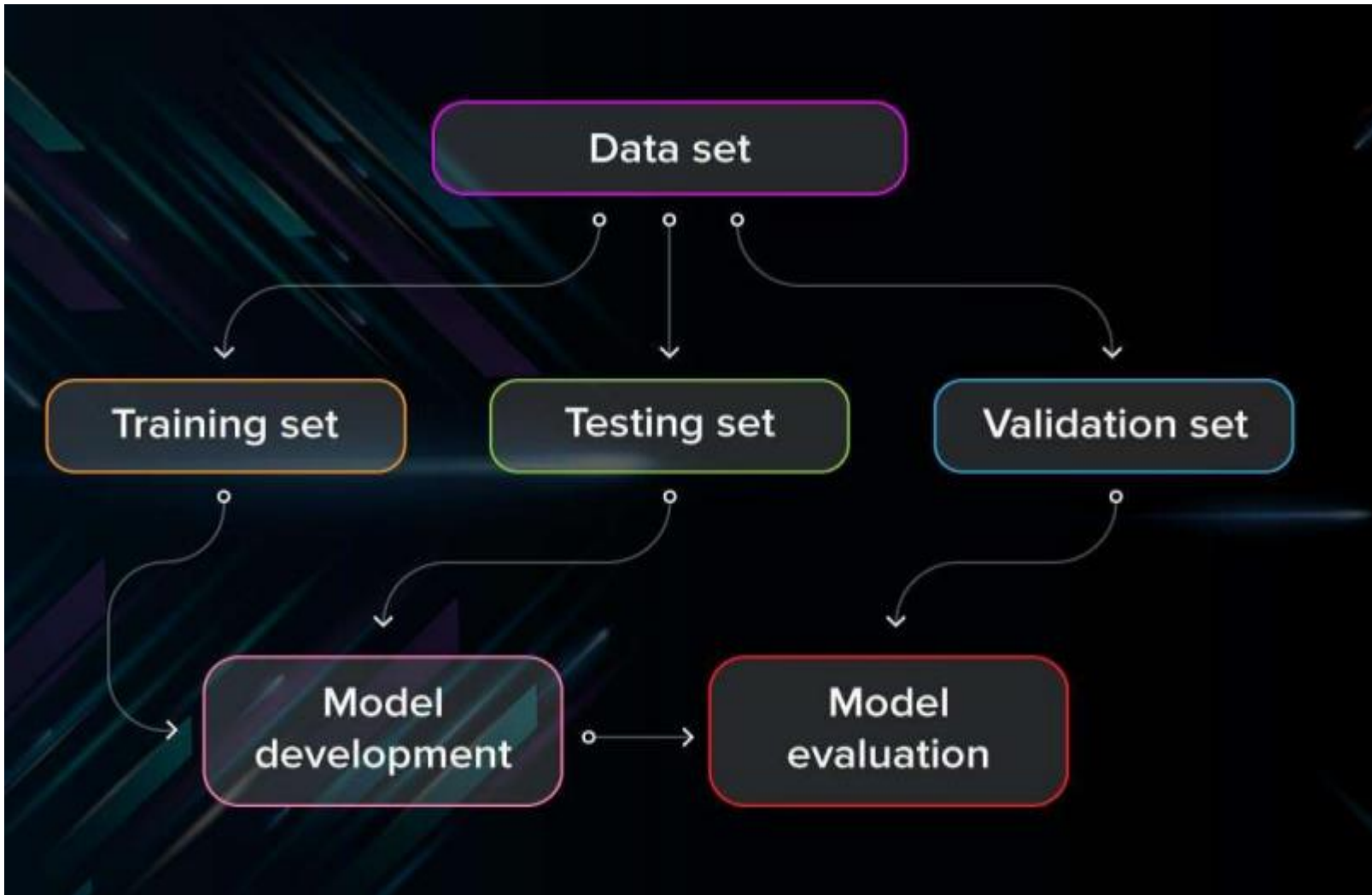


- Regression tests:
 - ✓ They cover already tested software to see if it doesn't suddenly break.
 - ✓ To Check whether your model breaks and test for previously encountered bugs.
- Integration tests:
 - ✓ This type of testing observes how multiple components of the program work together.
 - ✓ To Check whether the different components work with each other within your machine learning.



Training & Testing







- To use a training set to train the model. Then, to evaluate the performance of the model, you use two sets of data:

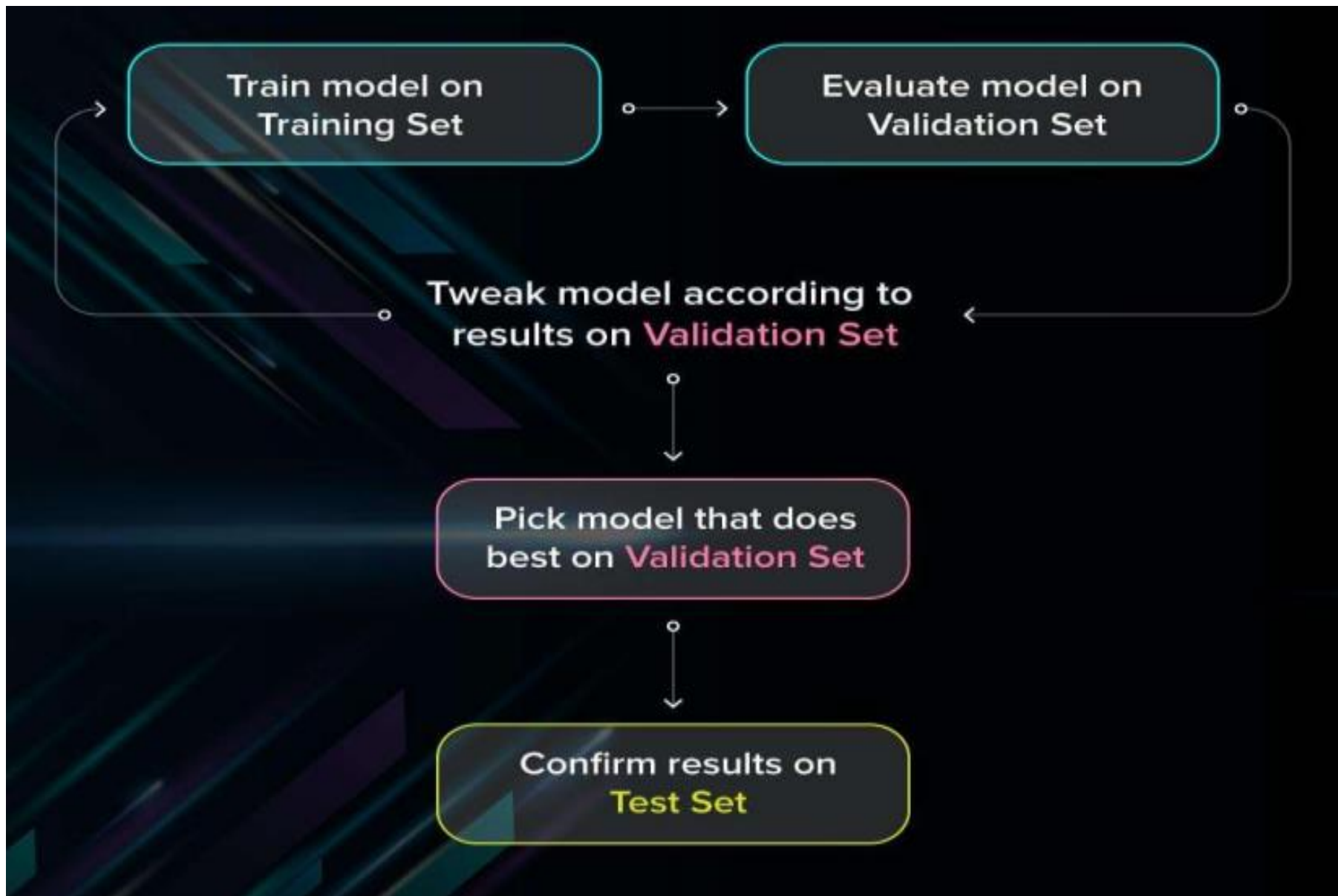
Validation set:

- Having only a training set and a testing set is not enough if you do many rounds of hyper parameter-tuning (which is always).
- you can select a small validation data set to evaluate a model.
- Only after you get maximum accuracy on the validation set, you make the testing set come into the game



Test set (or holdout set):

- Your model might fit the training dataset perfectly well.
- But where are the guarantees that it will do equally well in real-life?
- In order to assure that, you select samples for a testing set from your training set — examples that the machine hasn't seen before.
- Your test set should be large enough to provide statistically meaningful results and be representative of the data set as a whole.





- Tweak model: making minor adjustments to a model to improve its performance or results.
- Problems With Testing Machine Learning Models:
 - ✓ Quality assurance: whether the software works according to requirements
 - ✓ Identify defects and flaws during development and in production



challenges to testing machine learning models:

- Lack of transparency: Many models work like black boxes.
- Indeterminate modeling outcomes: Many models rely on stochastic(random) algorithms and do not produce the same model after (re)training.
- Generalizability: Models need to work consistently in circumstances other than their training environment.
- Unclear idea of coverage: There is no established way to express testing coverage for machine learning models.
- Resource need: Continuous testing of ML models is resource and time



Principles in Machine Learning Model Testing

- Testing is not easy, and testing machine learning models is even harder.
- Need to prepare your workflow for unexpected events while working with dynamic inputs, black-box models, and shifting input/output relationships.
- For this reason, some best practices in software testing:



- Test after introducing a new component, model, or data, and after model retraining
- Test before deployment and production
- Write tests to avoid recognized bugs in the future Testing machine learning models has additional requirements.

You also need to some testing principles specific to machine learning problems:

Robustness Interpretability Reproducibility