

Testing your ML pipelines

Applied ML days 2020

Presented by Kristina Georgieva





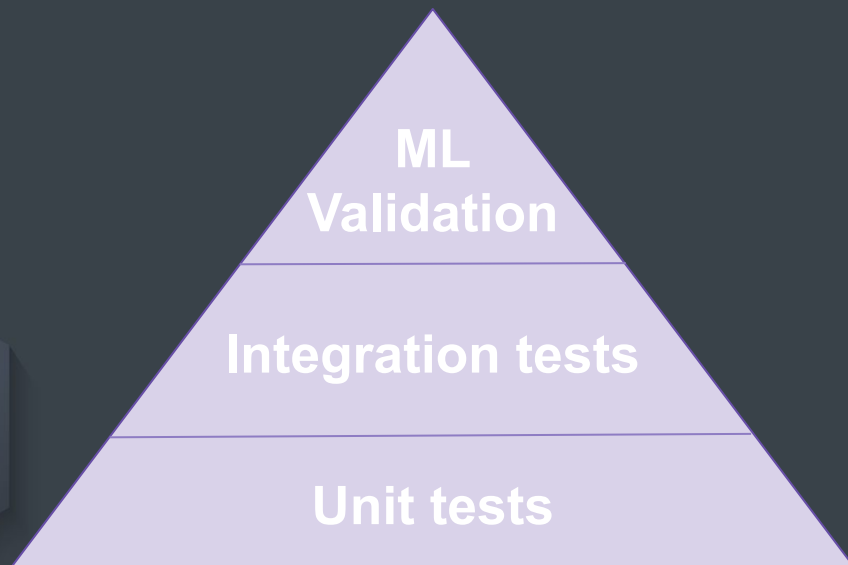
“You can’t test ‘machine learning’ because of its stochastic nature”

Random

“The engineer can do the production readiness things like testing”

Random

Testing pyramid



Unit testing

“It’s a system for testing your thoughts against the universe, and seeing whether they match”

Isaac Asimov.

Unit testing

Replacing /
Formatting
values

Products of
other features

Balancing
classes

Wrapping ML
libraries

Unit testing



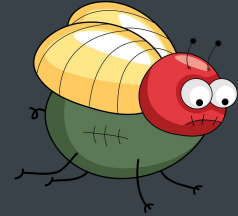
Ensures data is
cleaned as
expected



Helps split the
code



Documents the
cleaning
process



Helps
debugging

Integration testing

Because “The unclouded eye was better, no matter what it saw.”

Frank Herbert

Integration testing



Checks that
cleaning results
in expected
input to model



Can data be
consumed by
the model
config?



Helps refactor

Use case validation

Why? “To exhibit the perfect uselessness of knowing the answer to the wrong question.”

Ursula K. Le Guin.

Use Case Validation

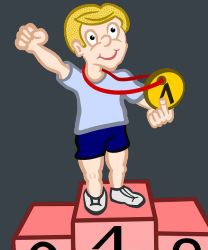
A model alone is not the product



Ensure that it helps solve the problem that the product needs



Documents business decisions around model



Can use to promote model



Provides visibility

In the wild: Monitoring

After all “Magic’s just science that we don’t understand yet”

Arthur C. Clarke.

Chat to me at the conference!



kristina.georgieva@bcgdv.com
@compArch1