



Università
della
Svizzera
italiana

Software Institute



P R E C R I M E

ASE 2021

DEEPMETIS: AUGMENTING A DEEP LEARNING TEST SET TO INCREASE ITS MUTATION SCORE



VINCENZO
RICCIO

@p1ndsvin



NARGIZ
HUMBATOVA

@nargiz_hmbtv



GUNEL
JAHANGIROVA

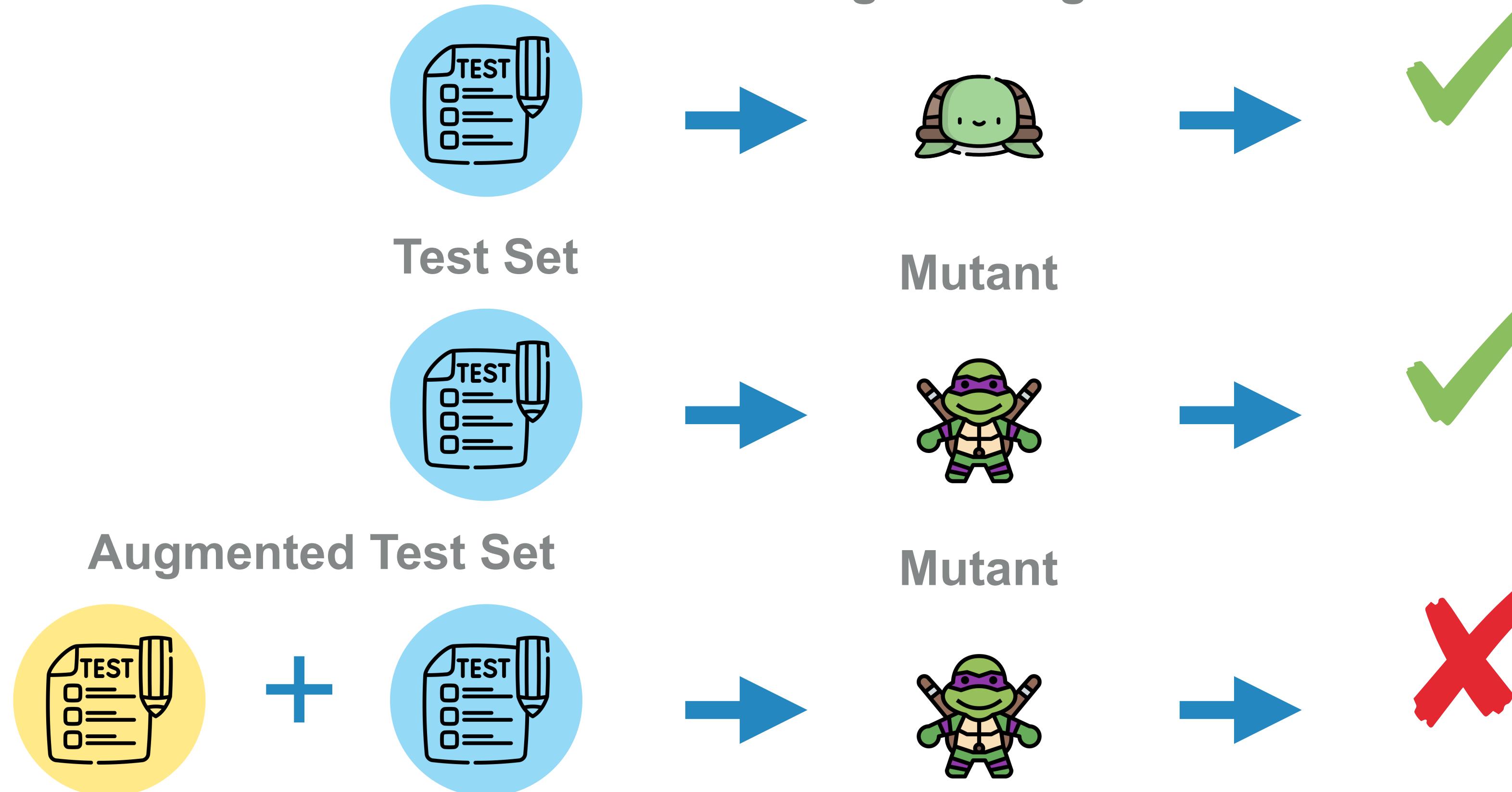
@guneli_j



PAOLO
TONELLA

@paolo_tonella

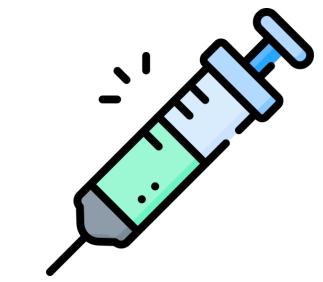
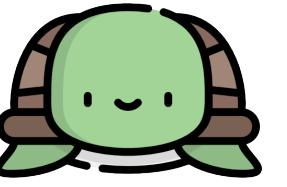
MUTATION TESTING



TRADITIONAL SOFTWARE MUTANTS

Original Program

```
public int calc(int x, int y){  
    int z = x + y;  
    return z % 3;  
}
```

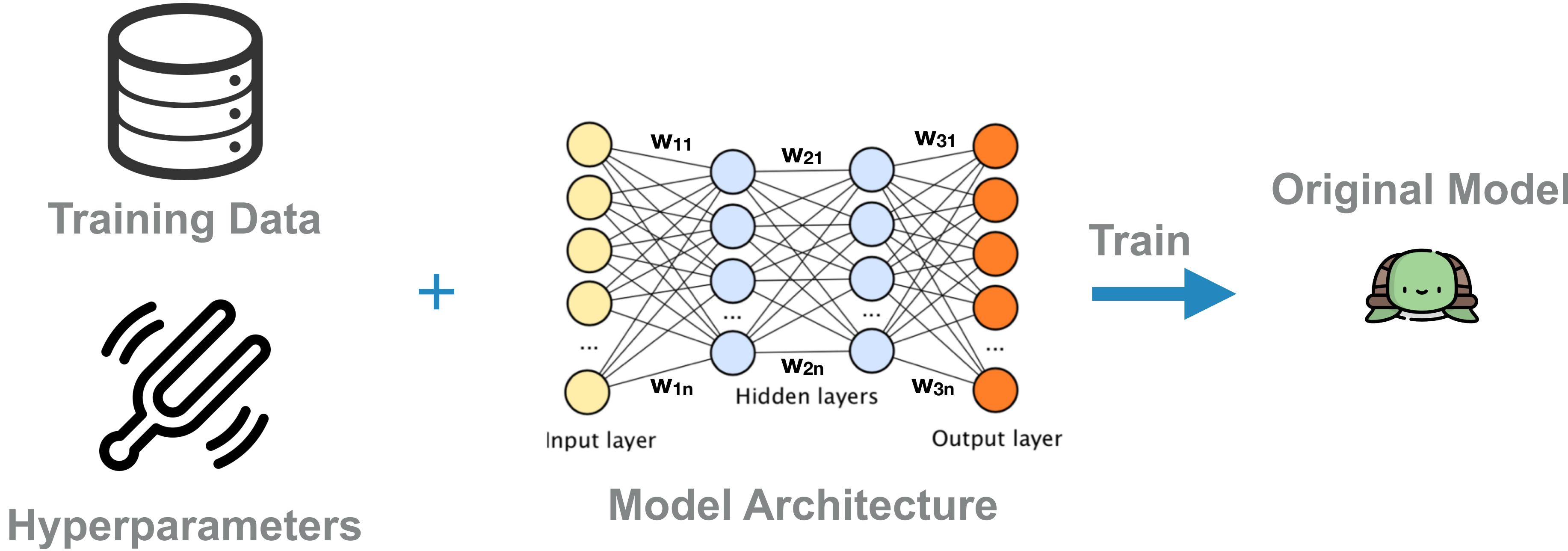


Mutant

```
public int calc(int x, int y){  
    int z = x + y;  
    return z + 3;  
}
```



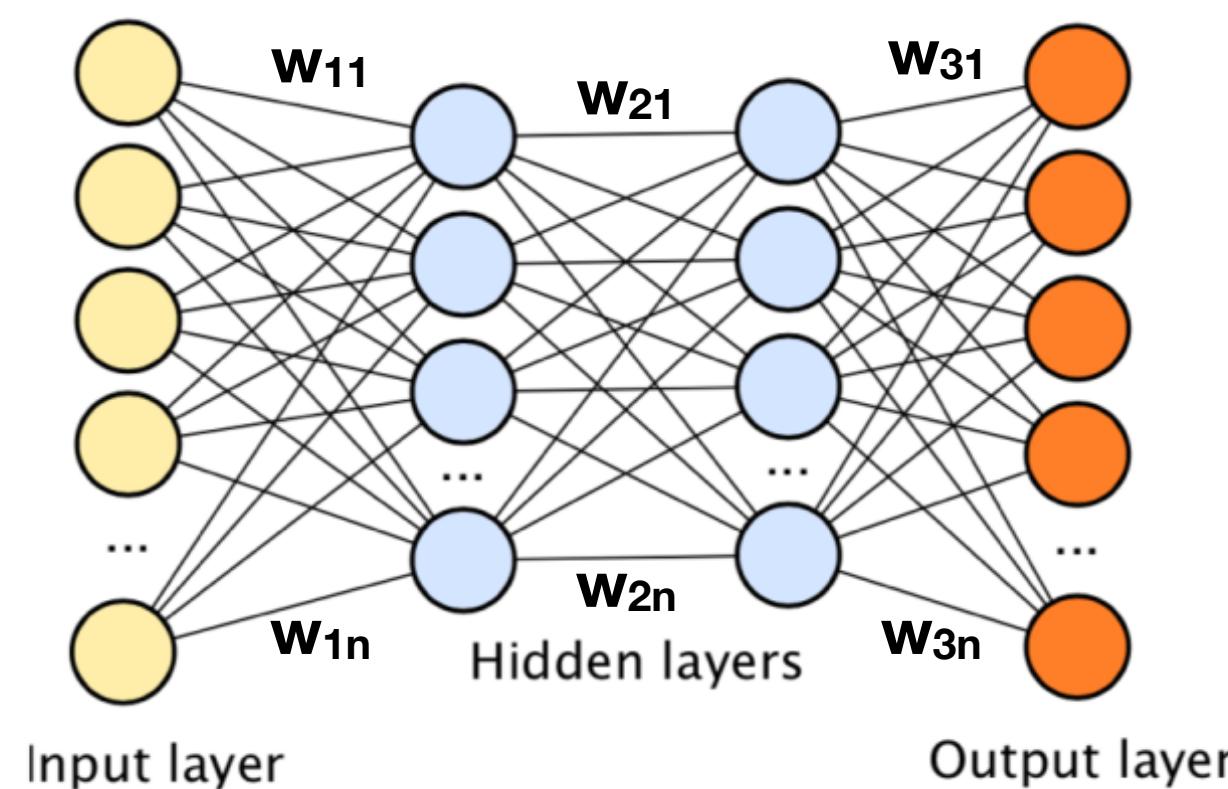
DL FAULT INJECTION WITH DEEPCRIME



DL FAULT INJECTION WITH DEEPCRIME



Hyperparameters



Model Architecture

Train
→

Mutant Model

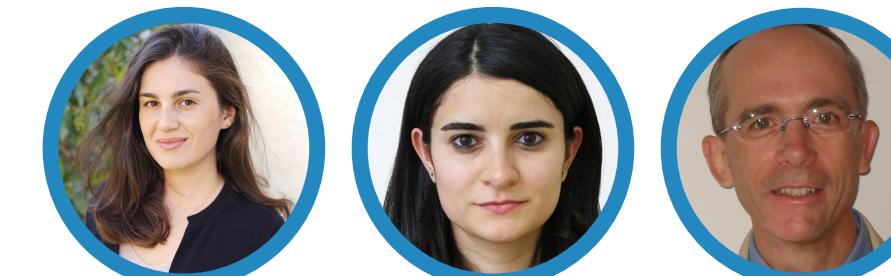


DeepCrime: Mutation Testing of Deep Learning Systems Based on Real Faults

Nargiz Humbatova
nargiz.humbatova@usi.ch
Università della Svizzera italiana (USI)
Lugano, Switzerland

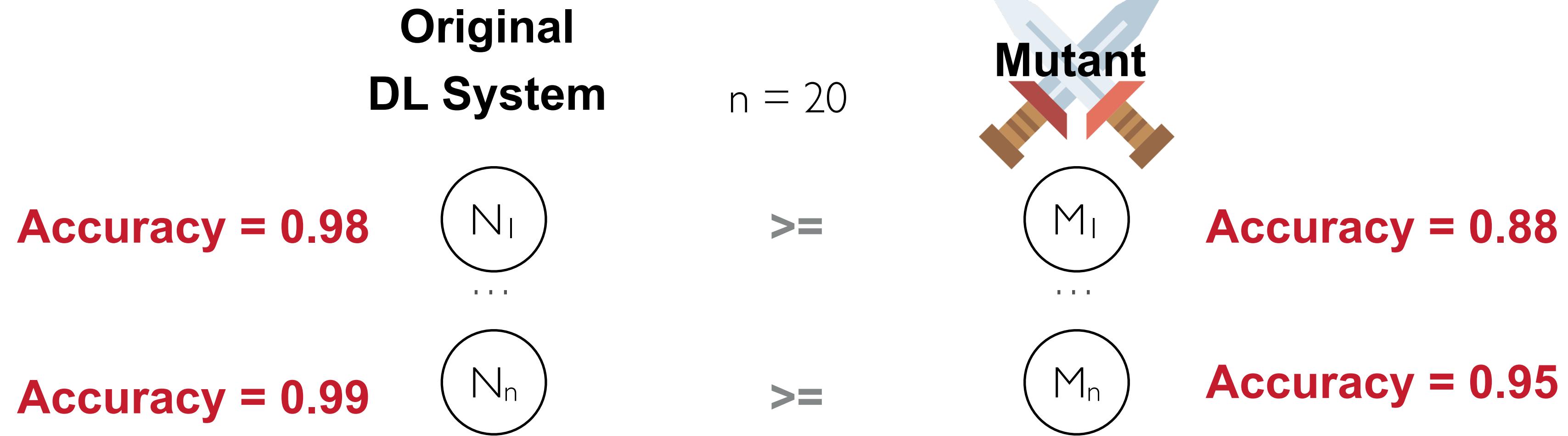
Gunel Jahangirova
gunel.jahangirova@usi.ch
Università della Svizzera italiana (USI)
Lugano, Switzerland

Paolo Tonella
paolo.tonella@usi.ch
Università della Svizzera italiana (USI)
Lugano, Switzerland



ISSTA 2021

STATISTICAL MUTATION KILLING



2020 IEEE 13th International Conference on Software Testing, Validation and Verification (ICST)

An Empirical Evaluation of Mutation Operators for Deep Learning Systems

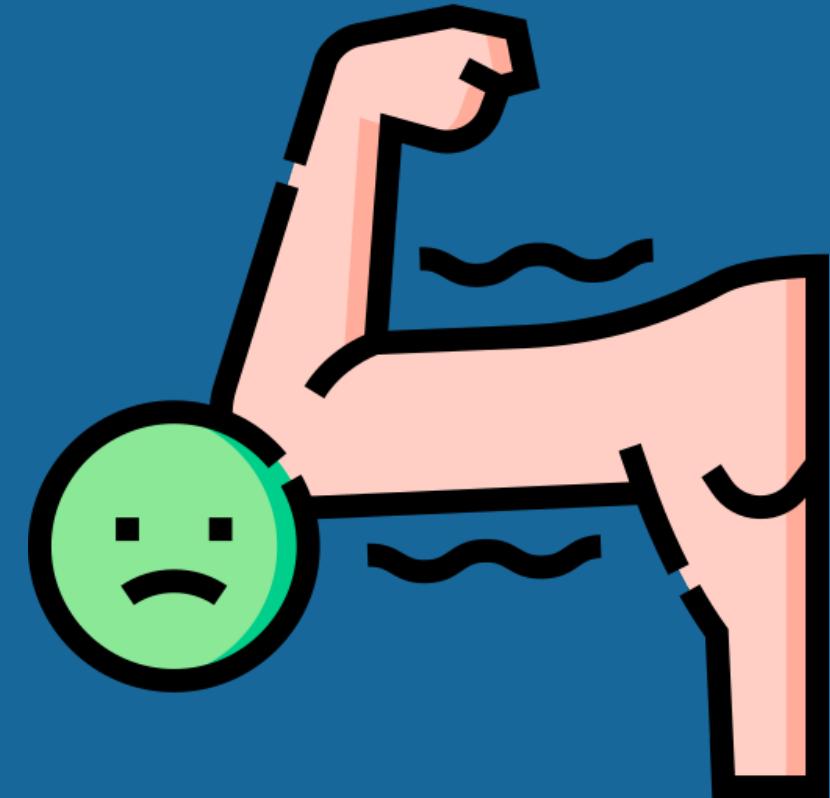
Gunel Jahangirova, Paolo Tonella
{gunel.jahangirova, paolo.tonella}@usi.ch
Software Institute - Università della Svizzera italiana, Lugano, Switzerland

*“A mutant is **killed** if the difference of the performance measurements across **n** re-trainings of the mutant and the original model are **statistically significant**”*



GOAL

WEAK TEST SET



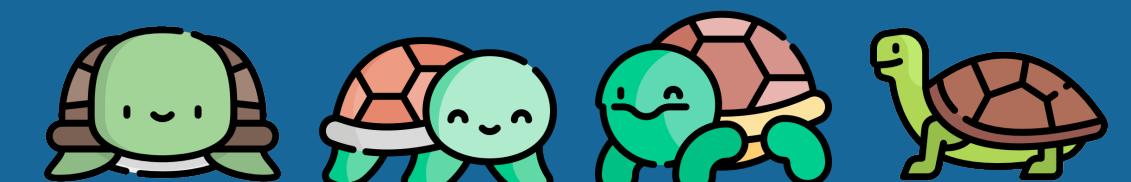
STRONGER TEST SET



MUTANTS



ORIGINALS



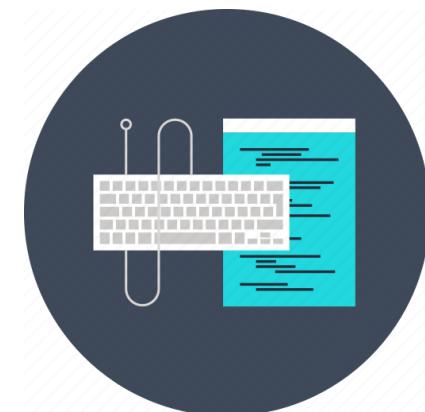
MUTANTS



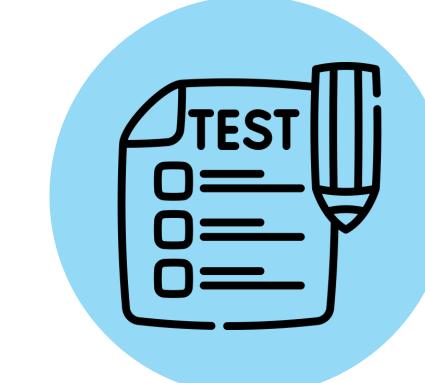
ORIGINALS

GOAL

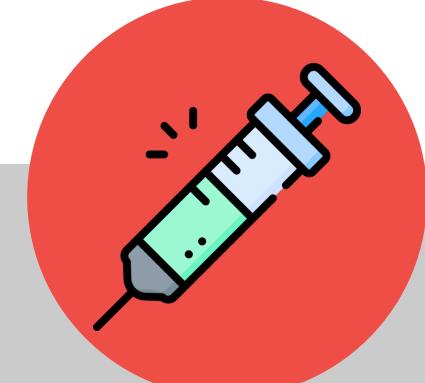
Original Program



Original Test Set

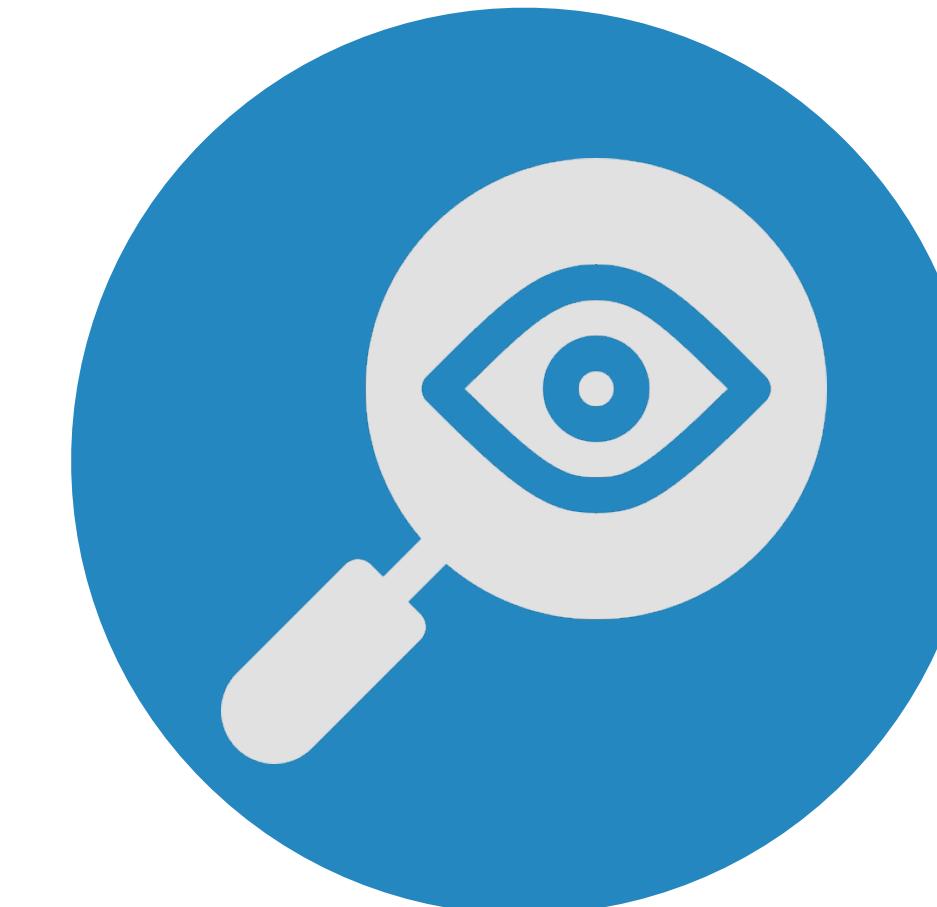


Mutation Operator

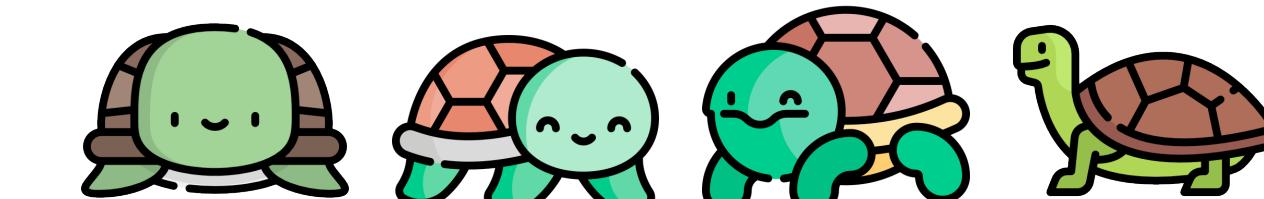


Remove portion of training data

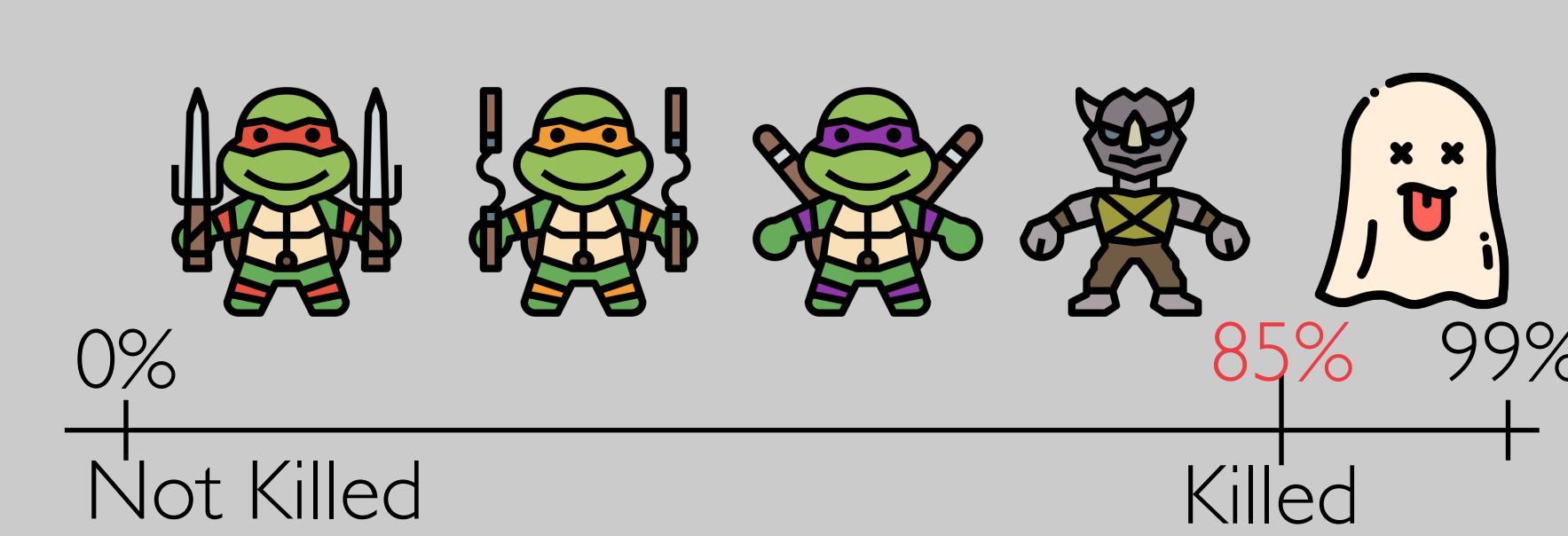
DeepCrime



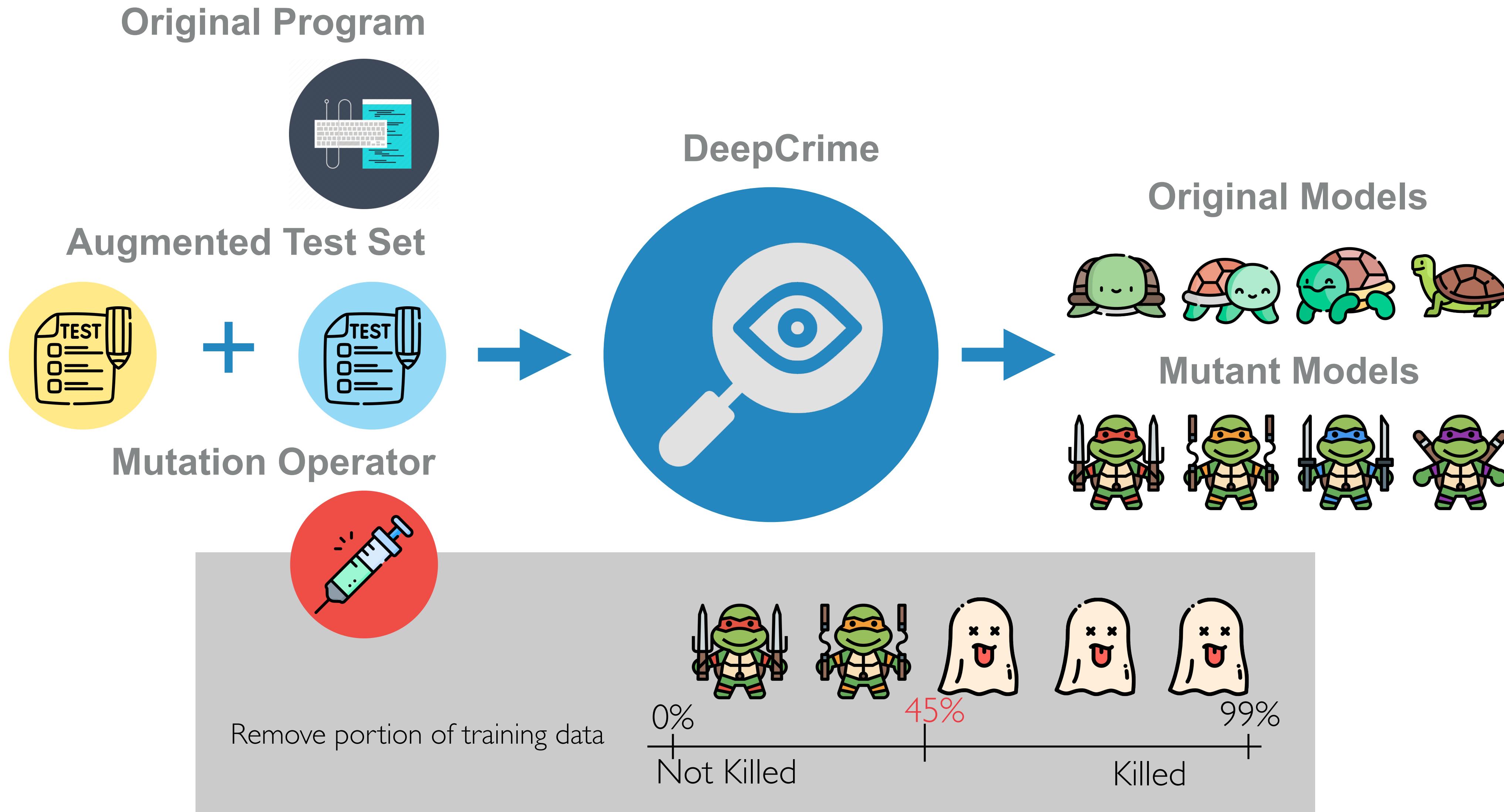
Original Models



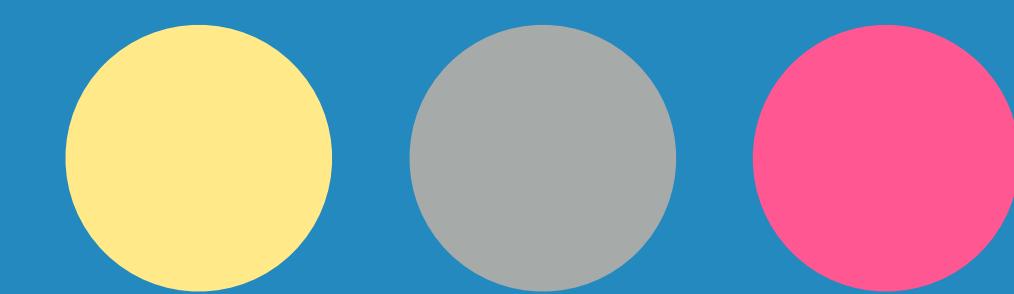
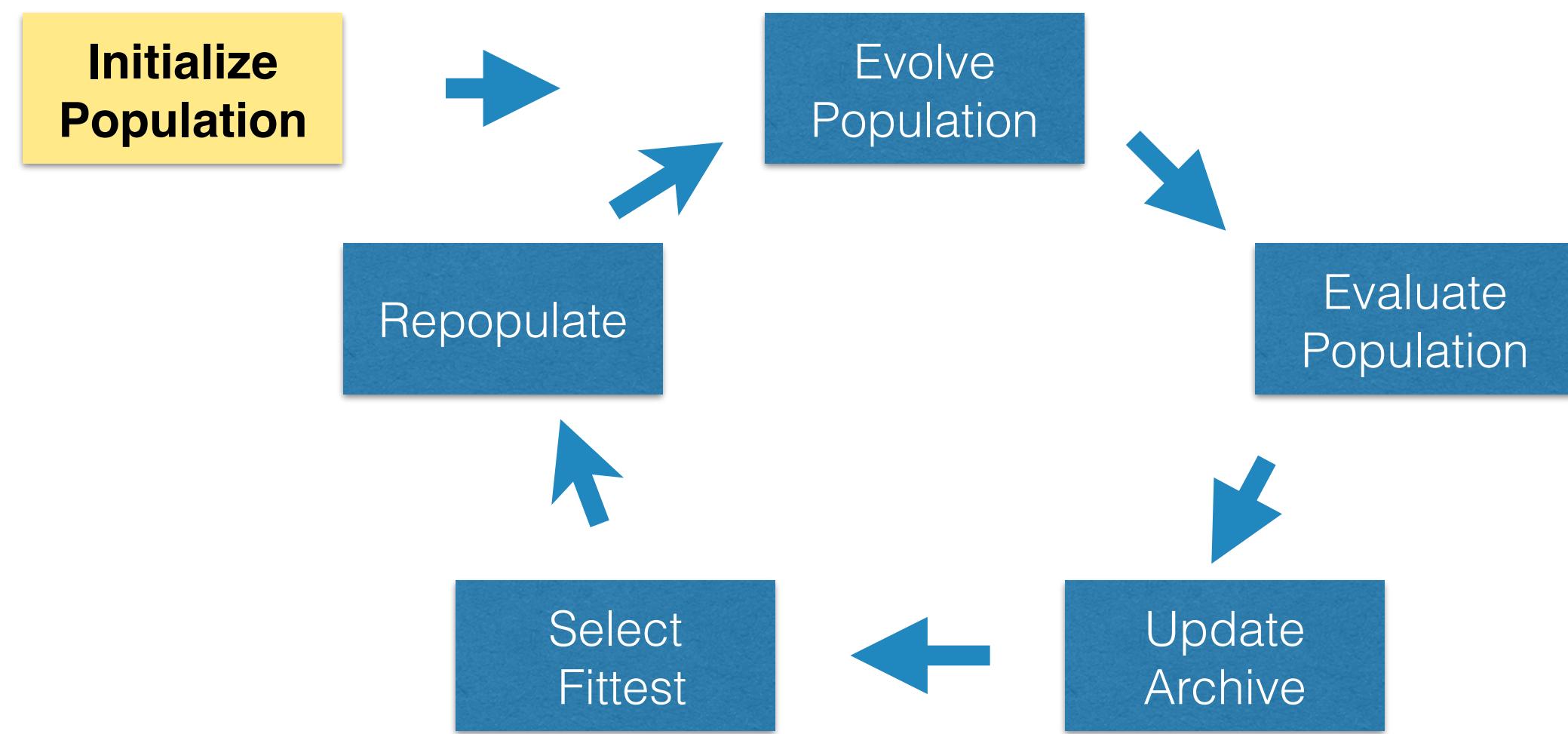
Mutant Models



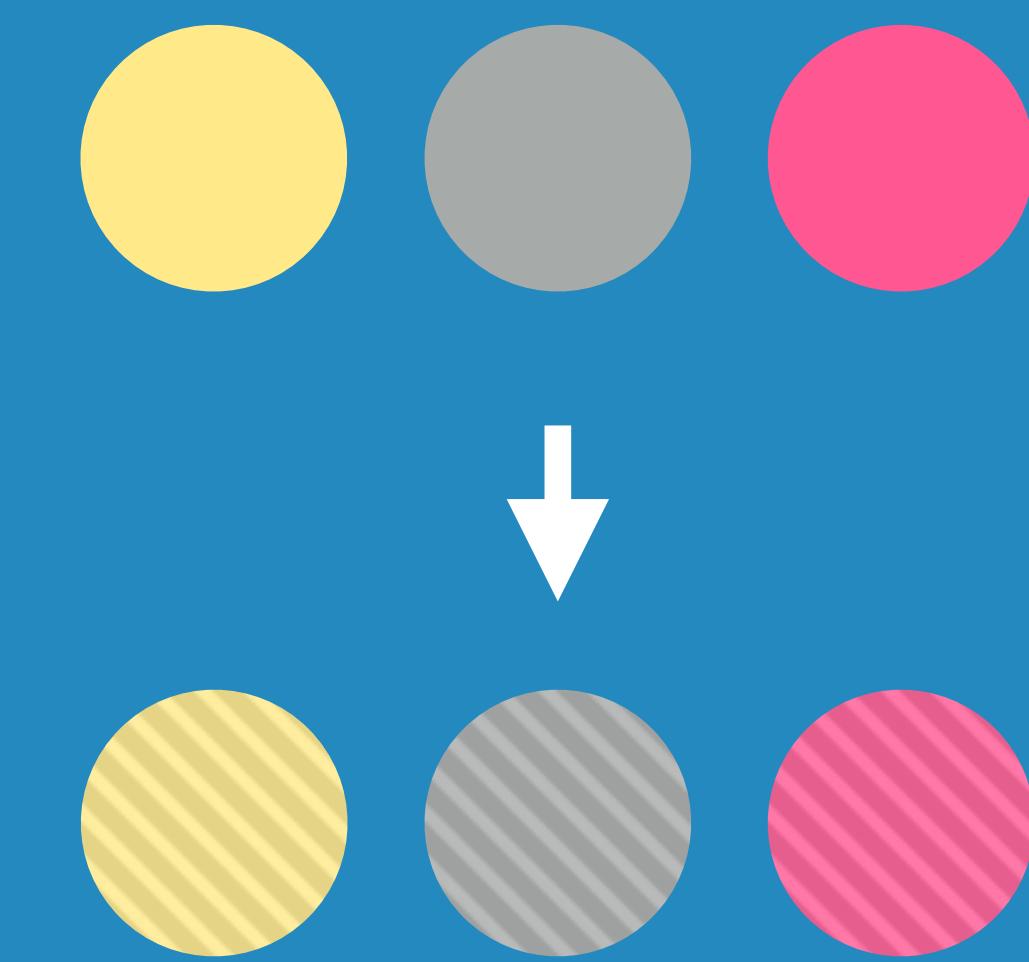
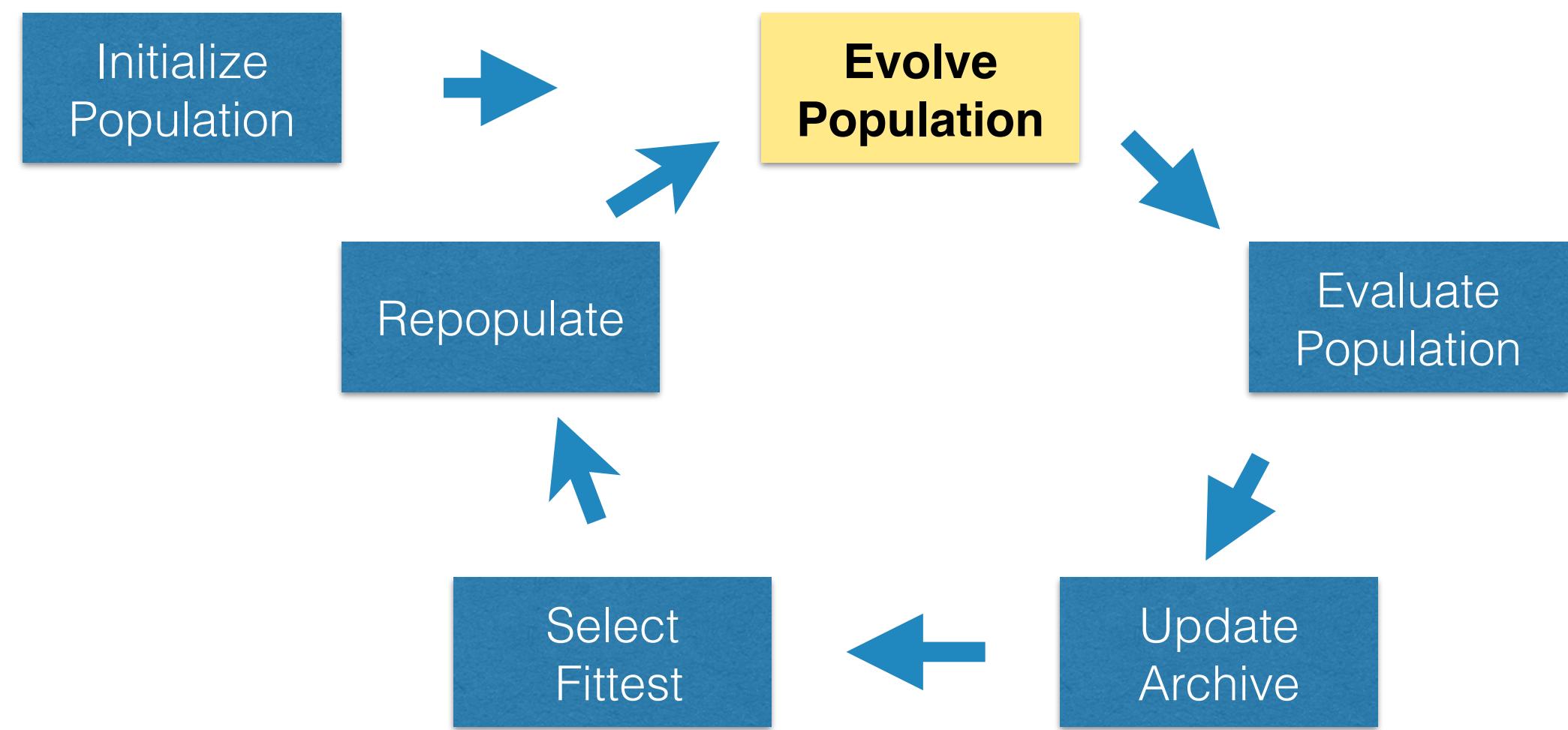
GOAL



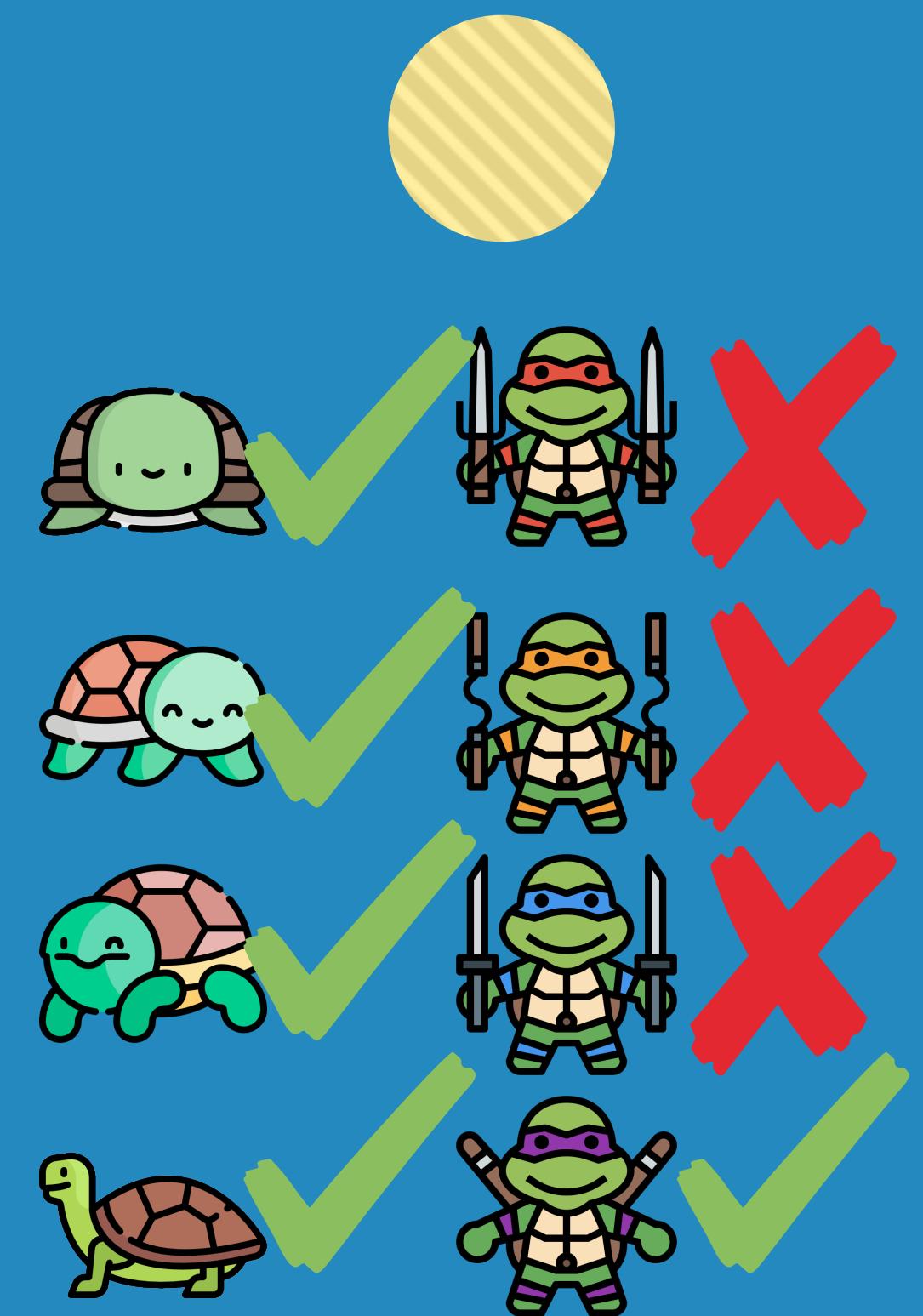
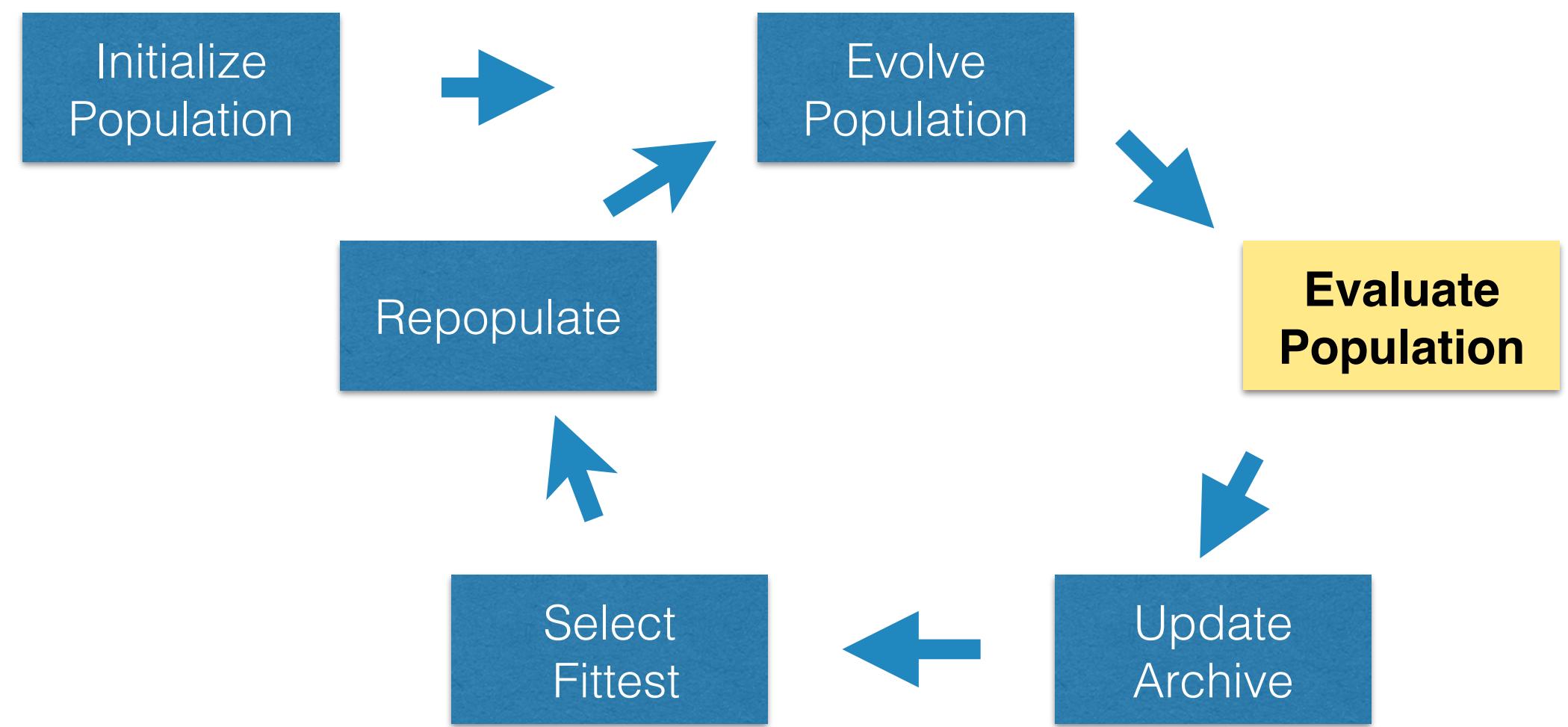
DEEPMETIS



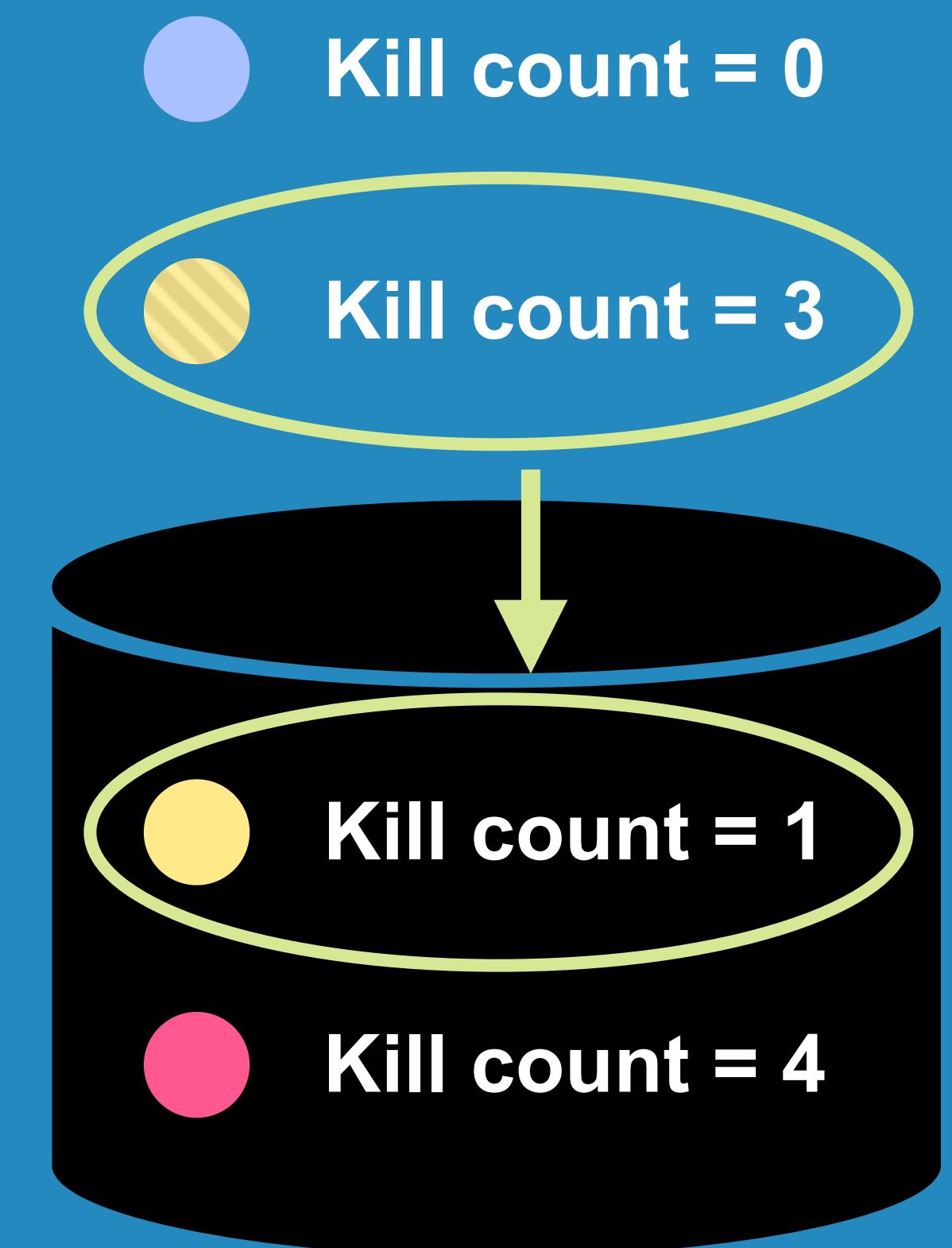
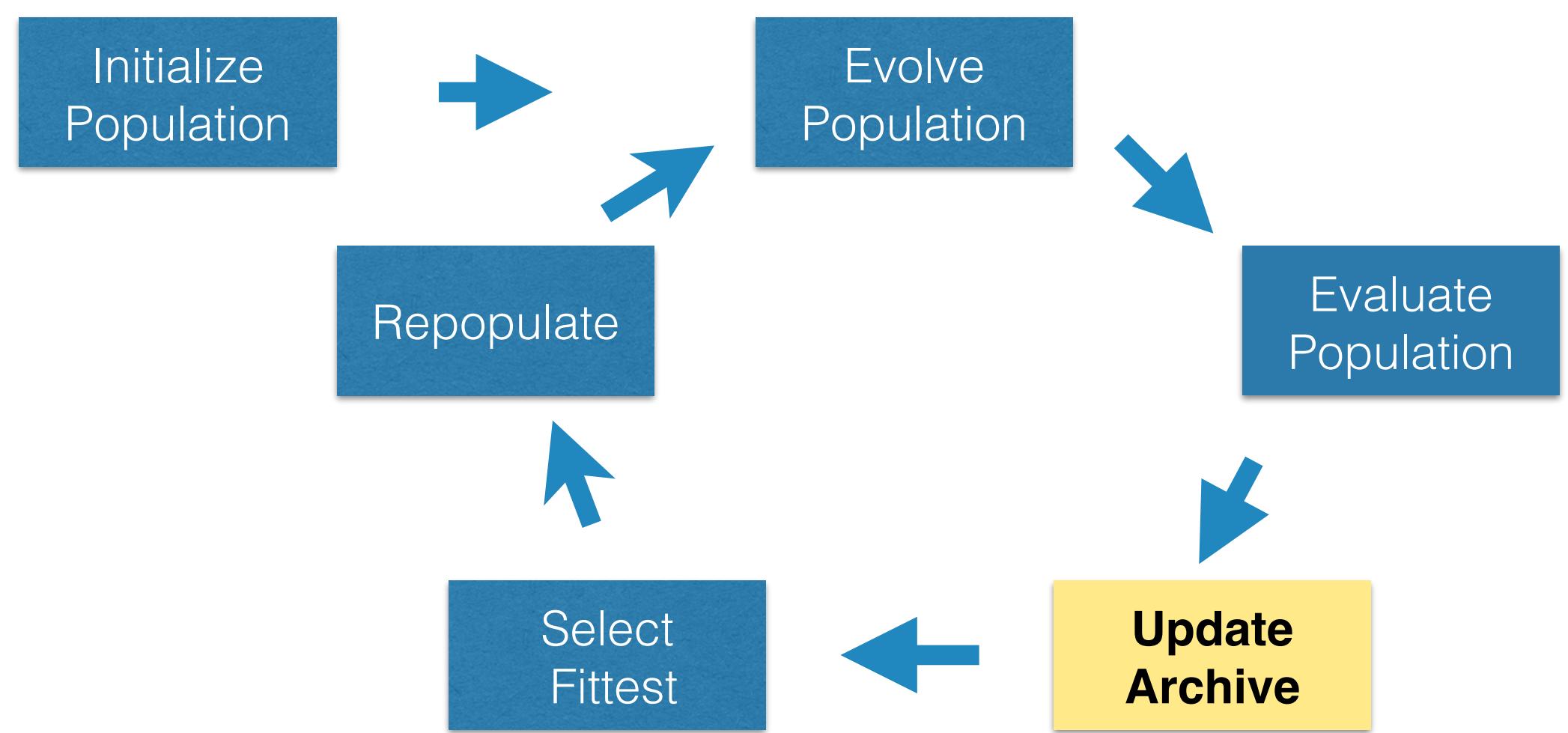
DEEPMETIS



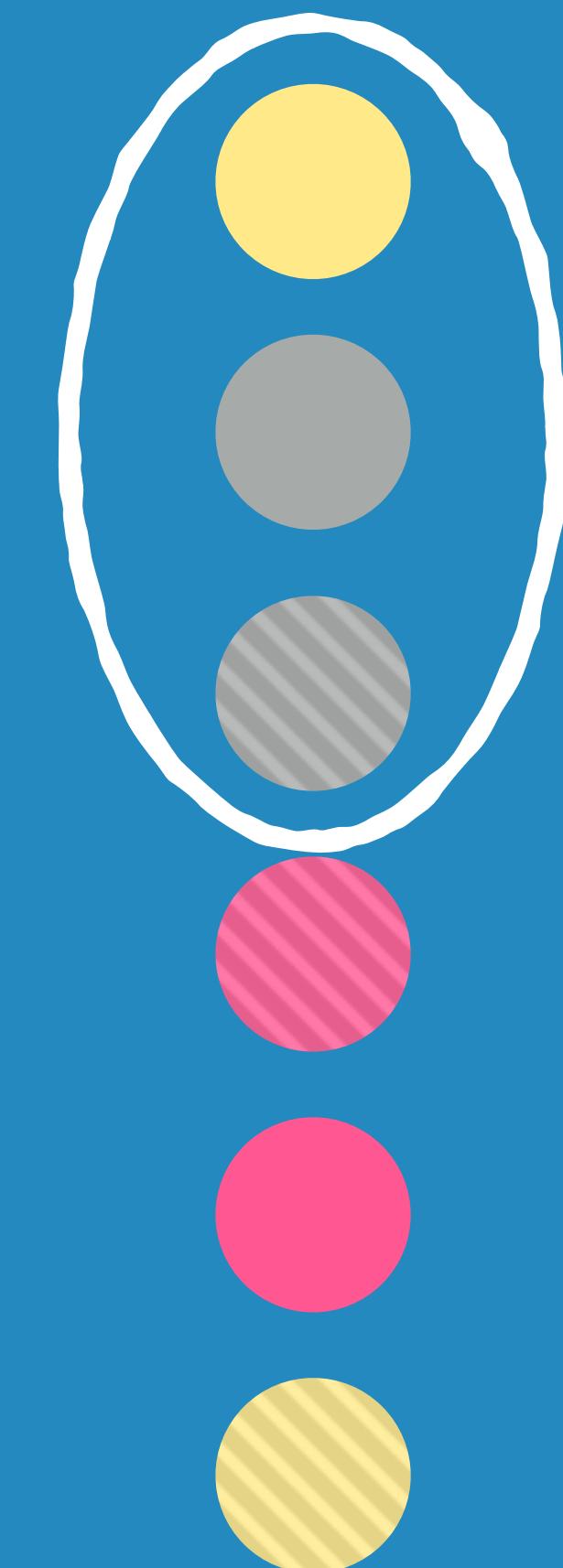
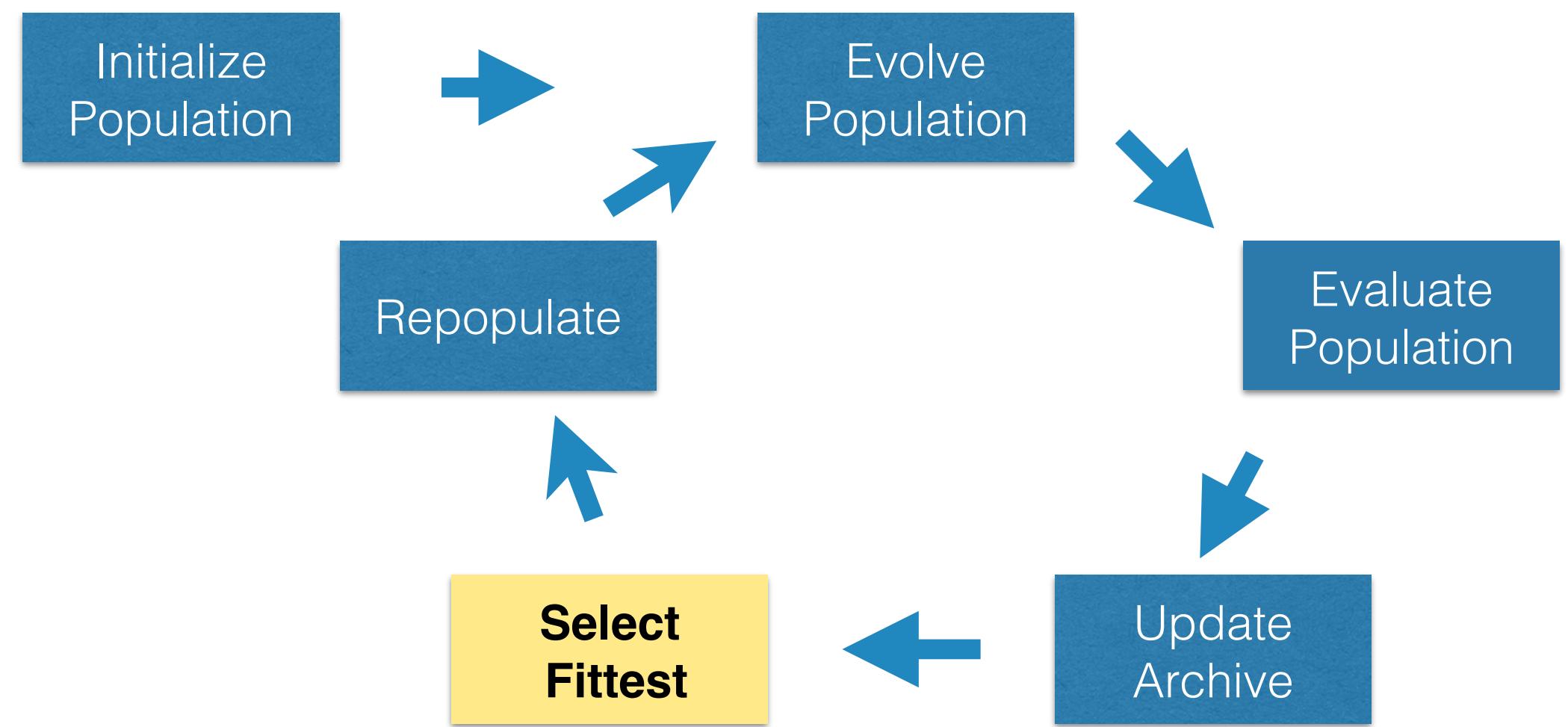
DEEPMETIS



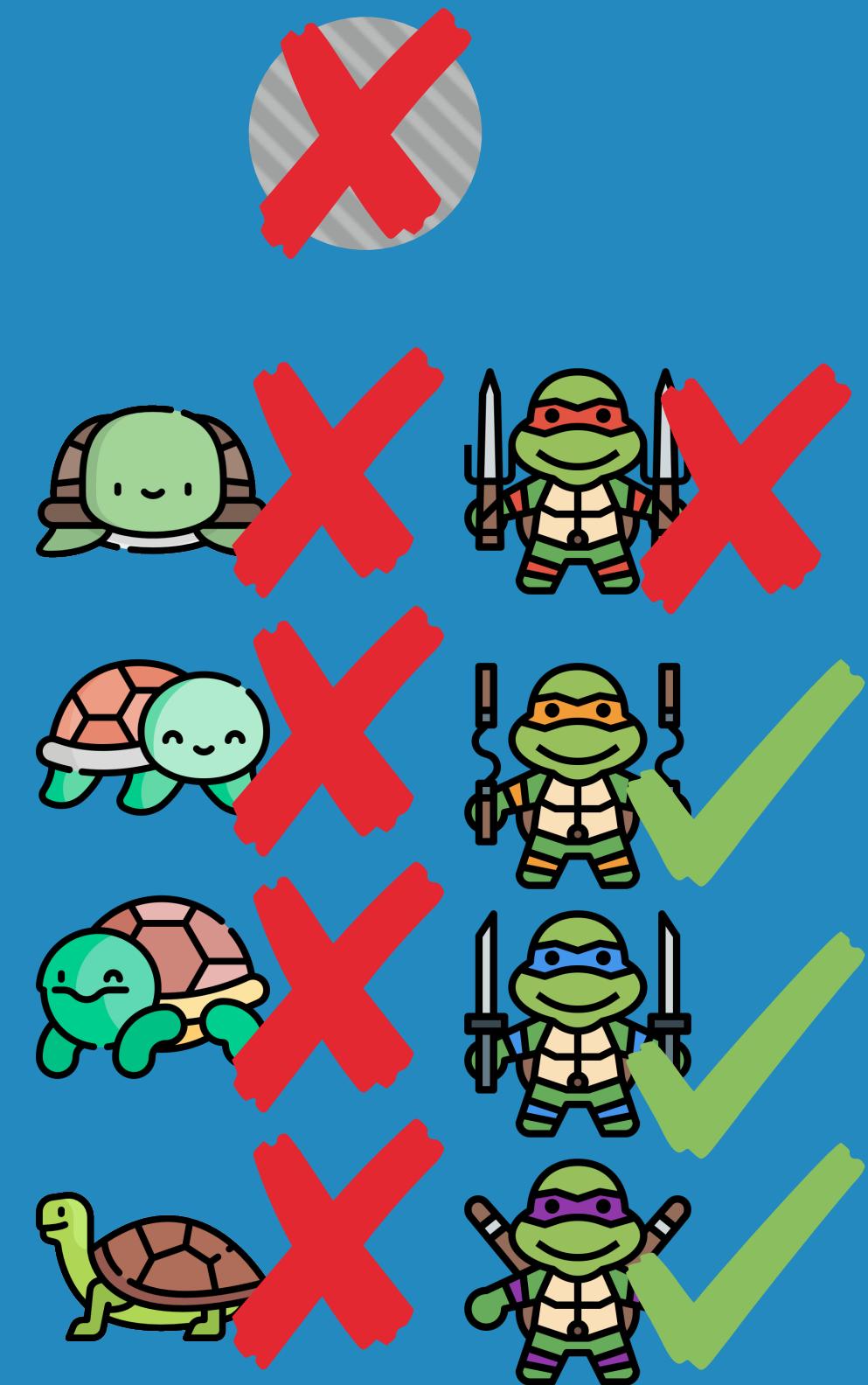
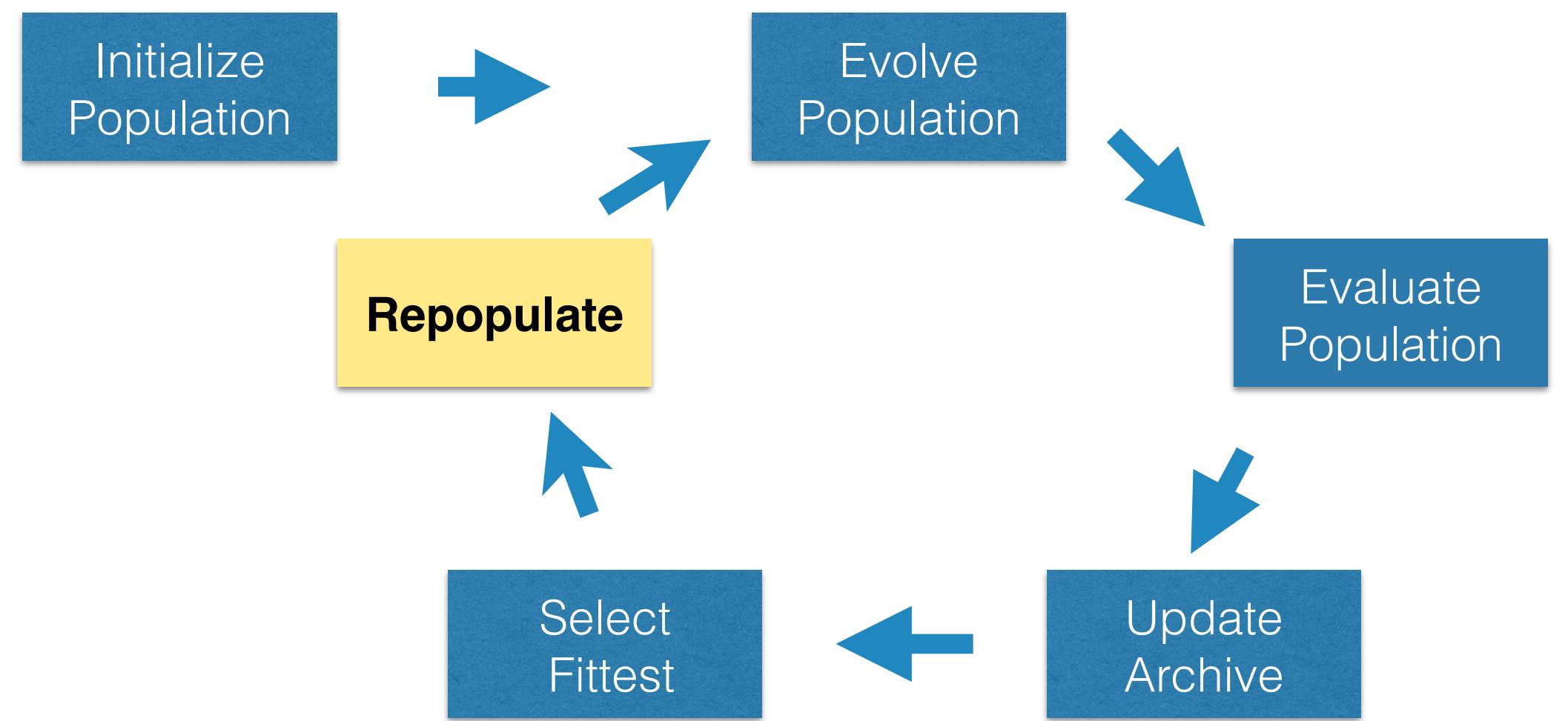
DEEPMETIS



DEEPMETIS

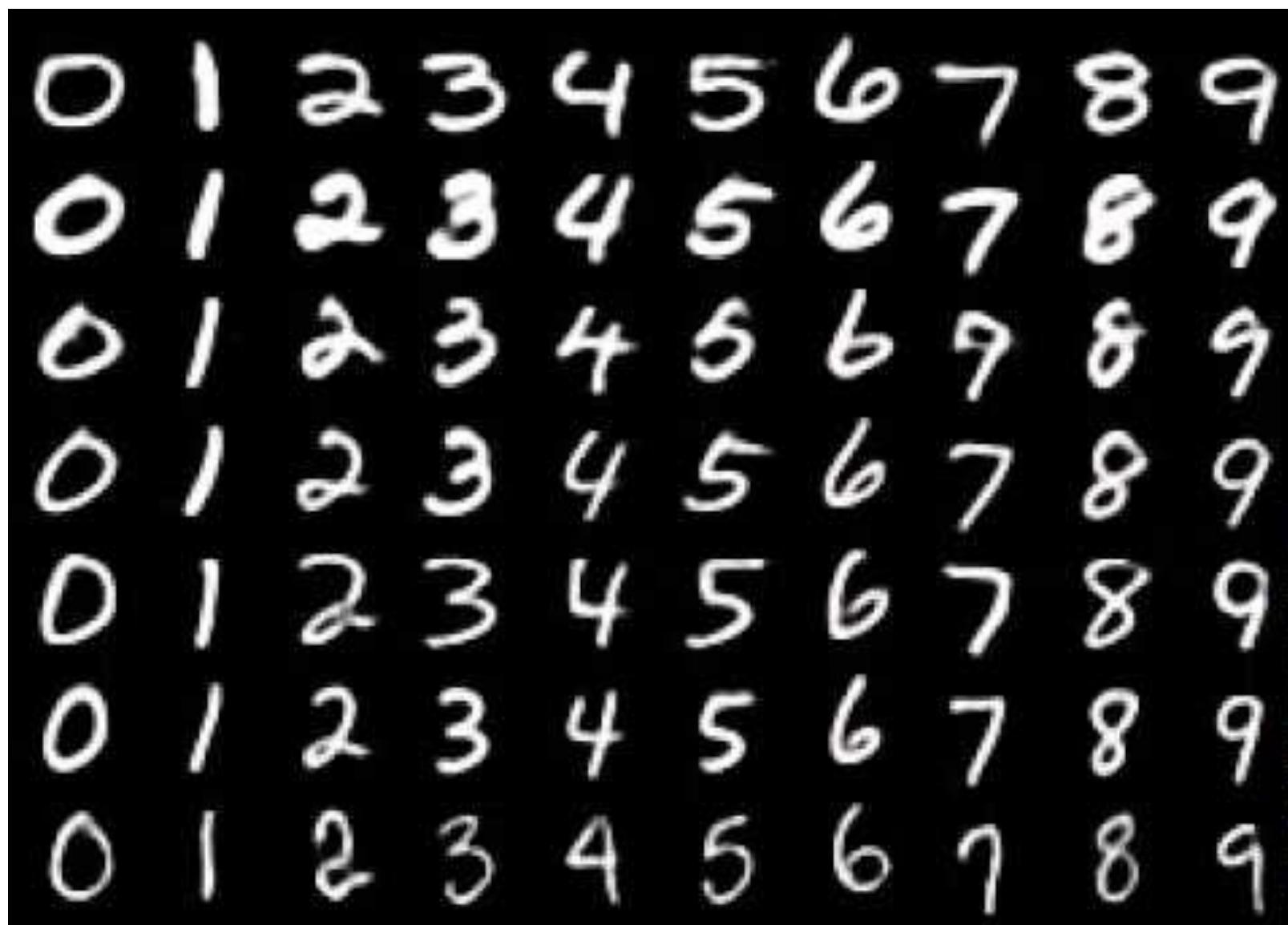


DEEPMETIS

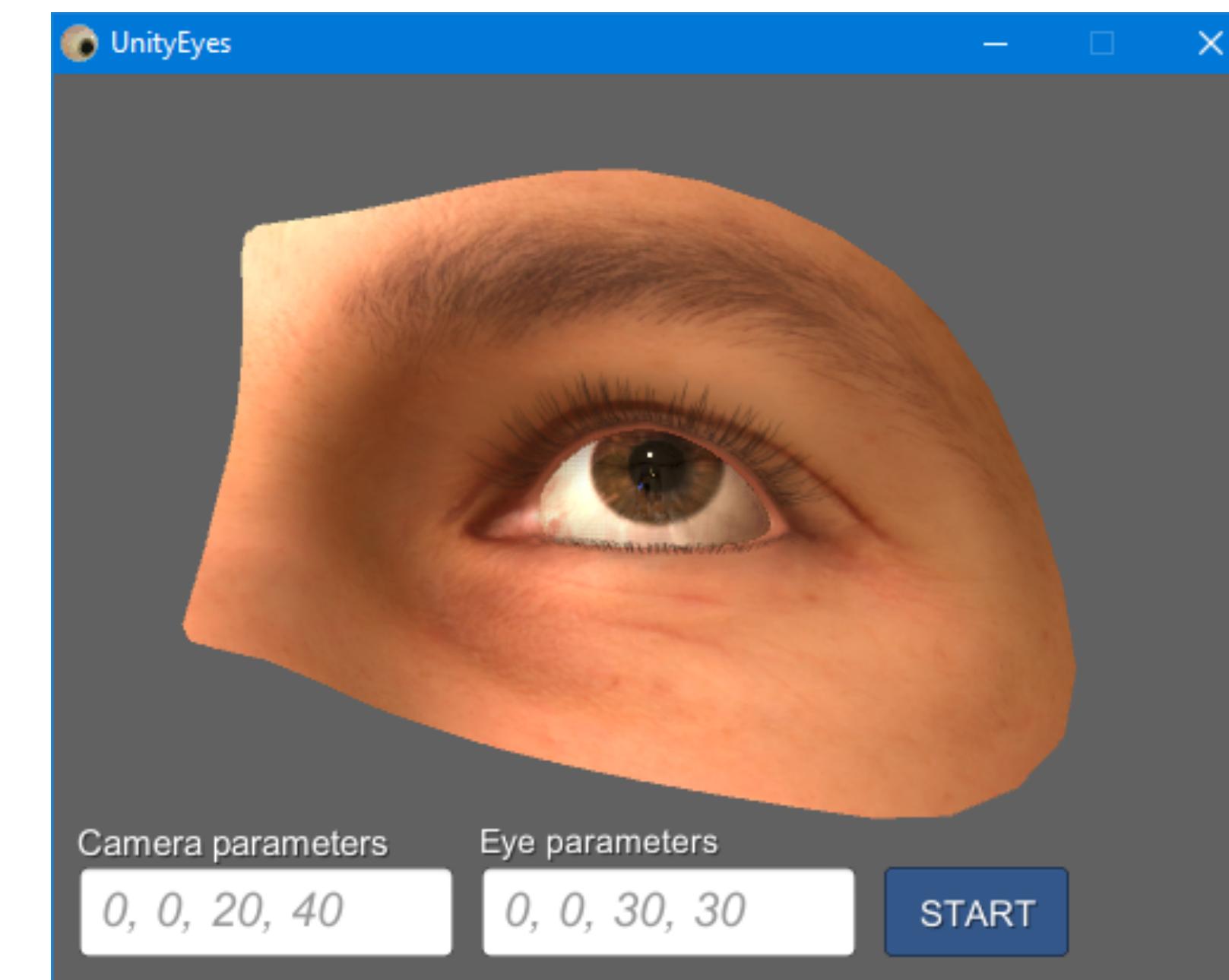


EXPERIMENTAL EVALUATION

MNIST



UNITYEYES



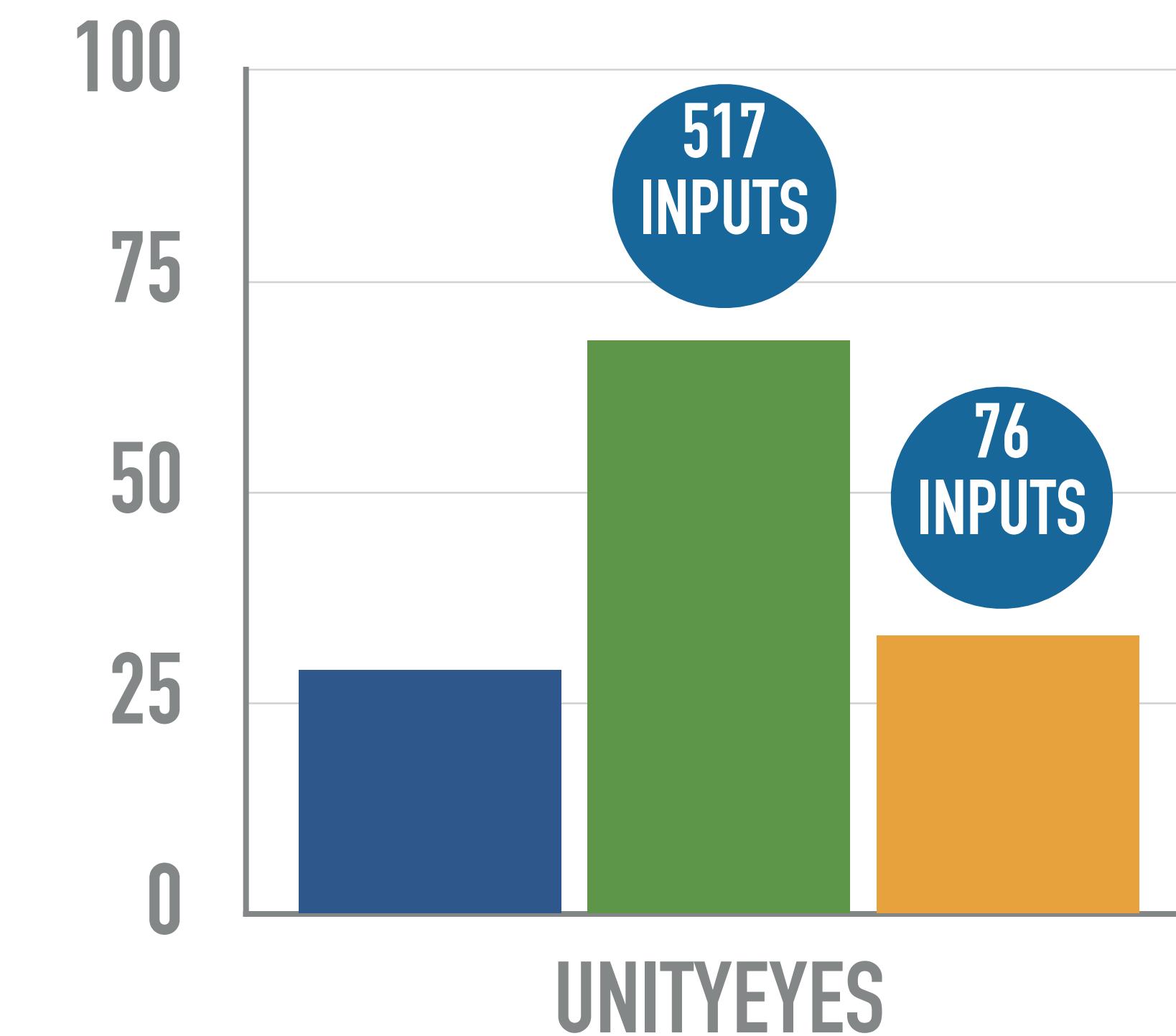
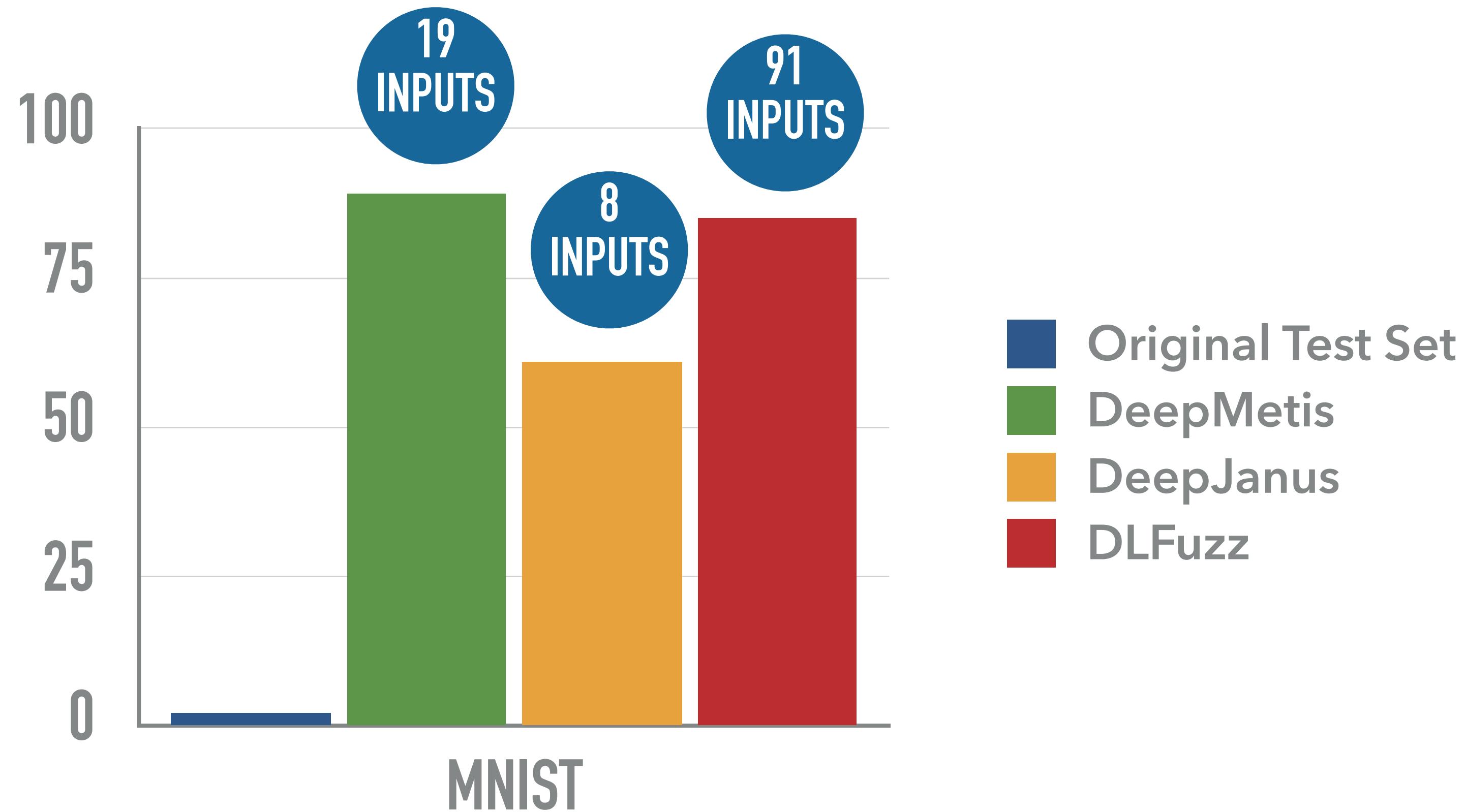
EFFECTIVENESS

GENERATION OF
MUTANT KILLING
INPUTS

COMPARISON
WITH STATE
OF THE ART

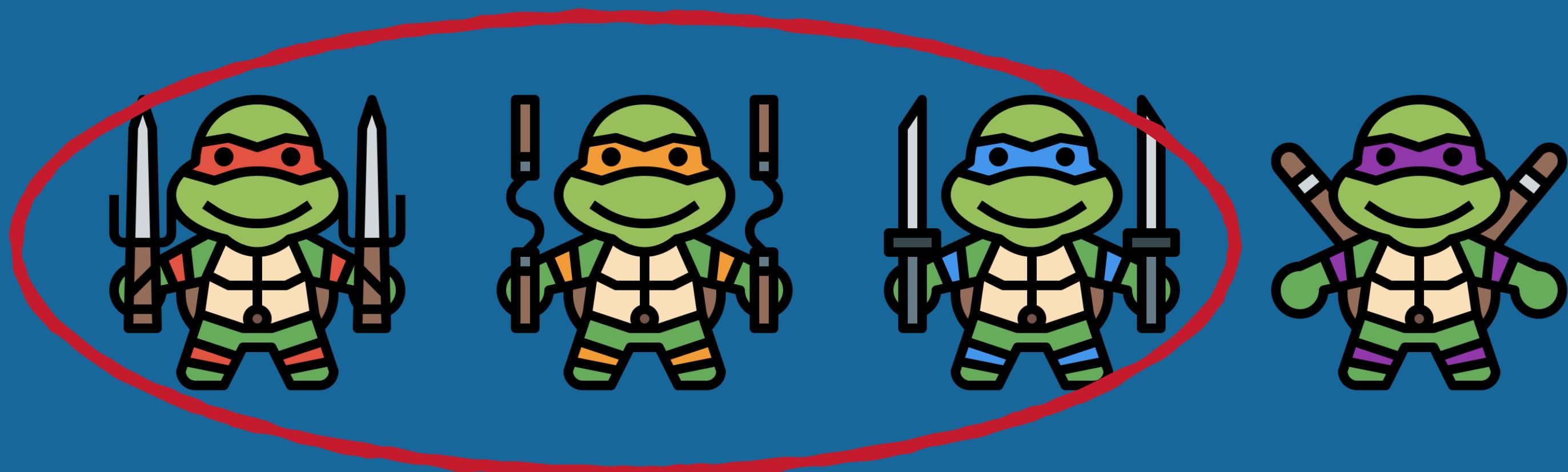
EFFECTIVENESS

Mutation Killing (%)

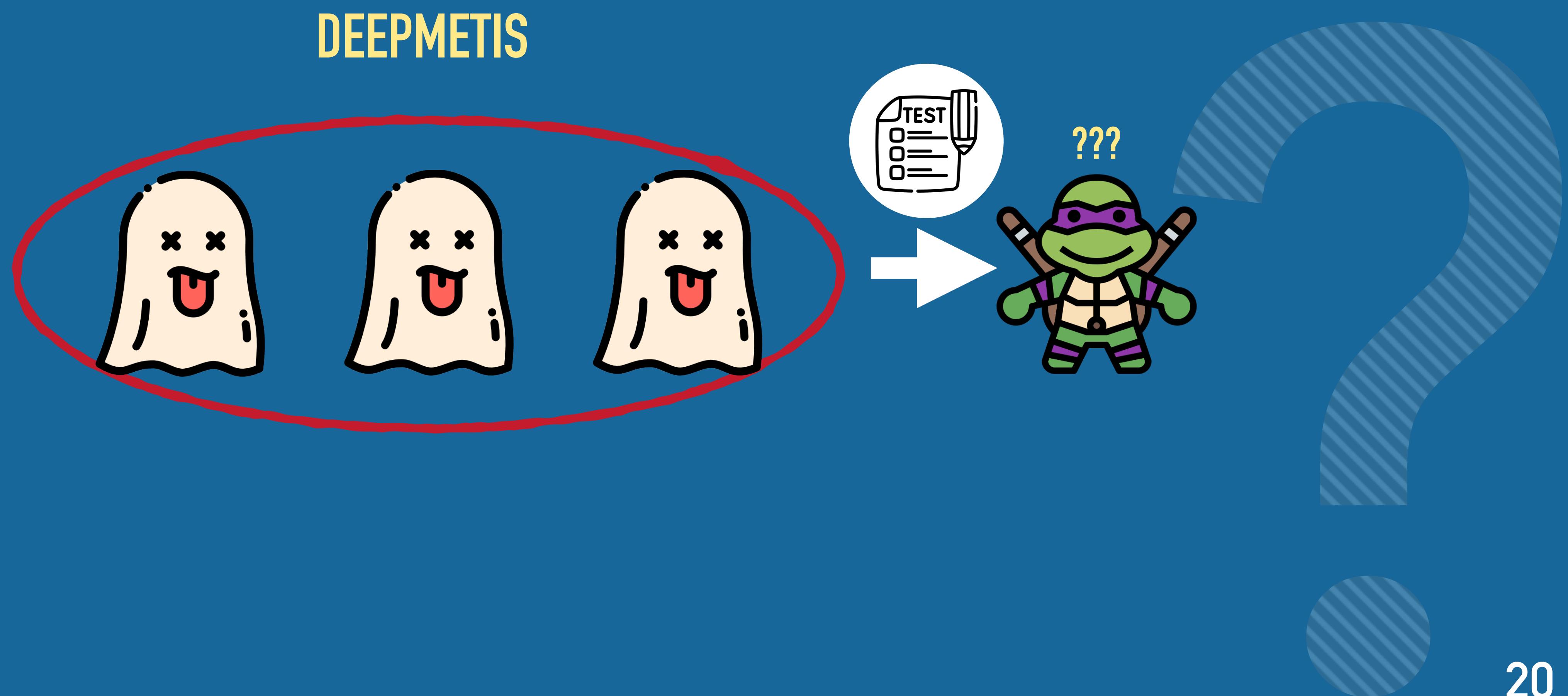


FAULT DETECTION

DEEPMETIS



FAULT DETECTION



FAULT DETECTION

