ML4Science

Week 5

What we did

- implemented PINN with parallel branching
- gathered final setups and results (see branch “combination”)
- started writing the report (introduction, physics background, data and models and methods)

Last week recap: our model fails on slowly evolving functions

All simulations

“Normal”, 80% of training data

Upper bound in range [0.1, 1)

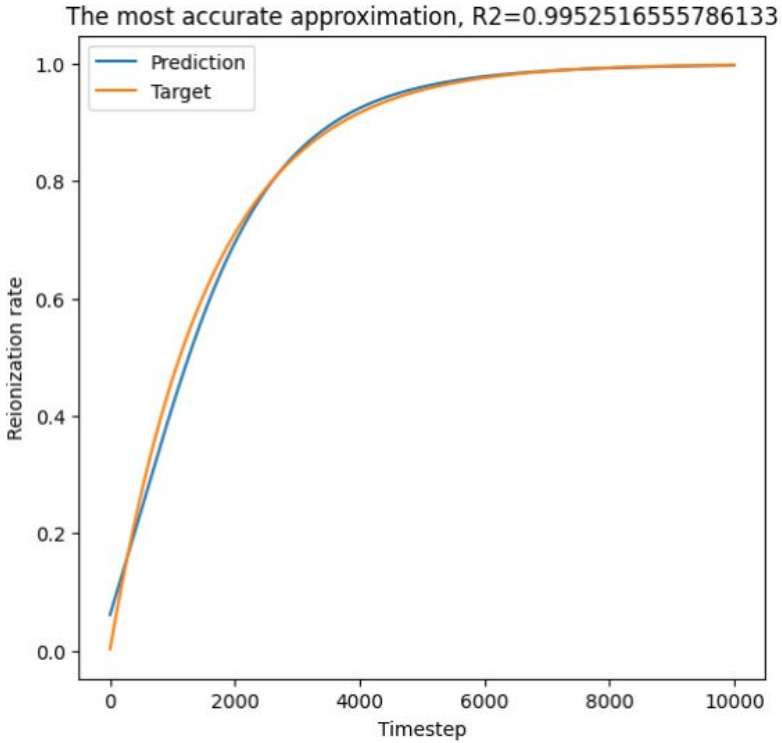
High and stable quality of approximation

“Out-of-distribution”, 20% of training data

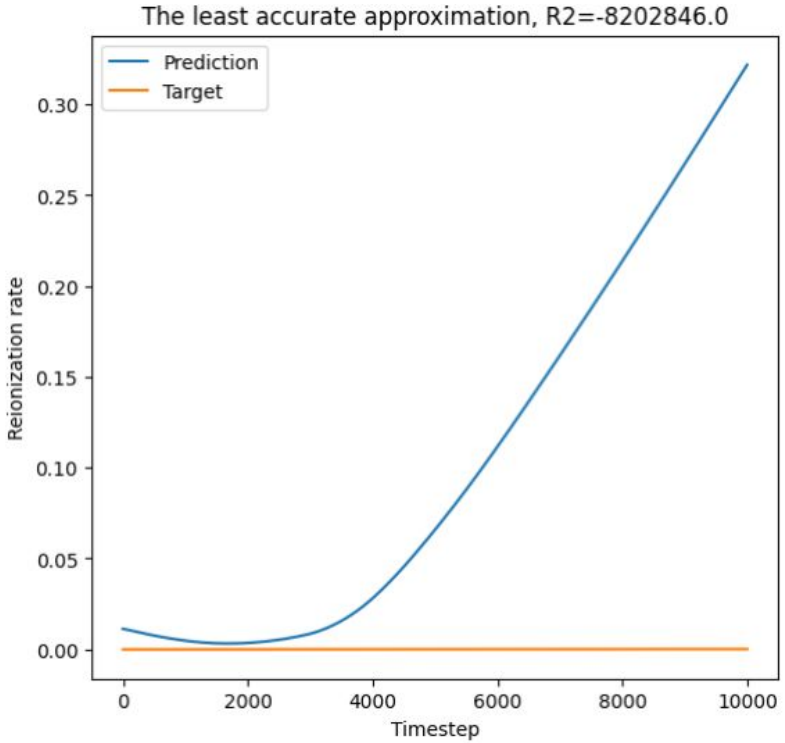
Upper bound in range (0, 0.1)

Unsatisfactory quality of approximation

Unbalanced data result in unstable quality of our model



“Normal” simulation



“OOD” simulation

Boundary investigation for physics parameters

We train a basic model on a single “normal “simulation and then make predictions for the rest of data

We measure the range of physics parameters, in which model doesn’t fail

Criterion: R2 score for a simulation > 0.99

Results: Gamma = (1.7 ± 0.2)e-13

Contains small portion of values in the training set

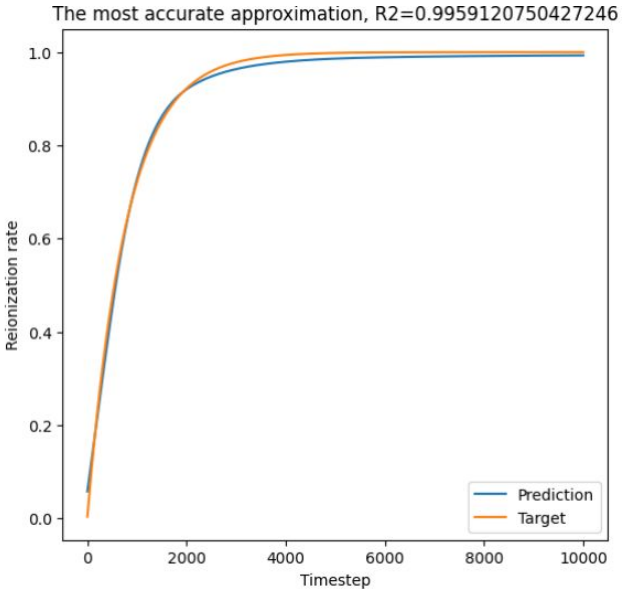
nH = (10 ± 8)e-8

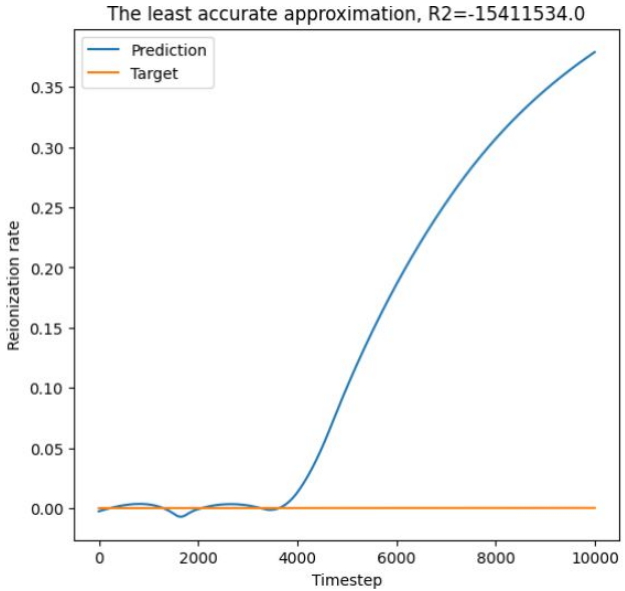
Contains most of values in the training set

Proposed modifications of the training pipeline

1. **Re-sampling idea:** let’s balance the low frequency of “ood” simulations by sampling them with a higher probability during training
2. **Scaling:** sensitivity of the model comes from Gamma value, which is small compared to concentration and timesteps values
3. **Parallel model configuration**: let’s process timesteps and physics parameters by separate branches

Before modifications

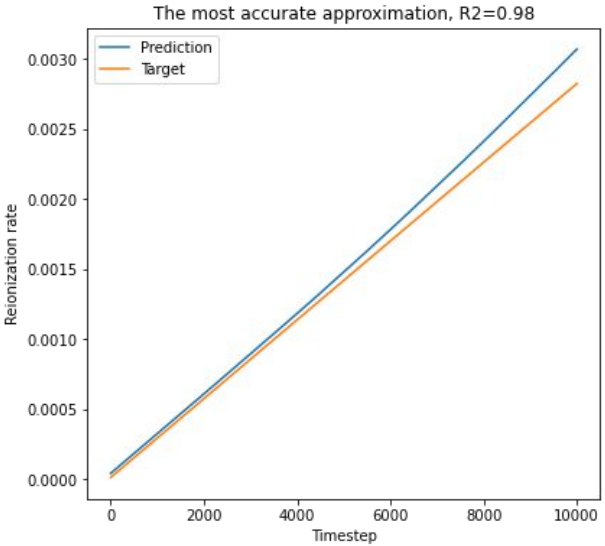


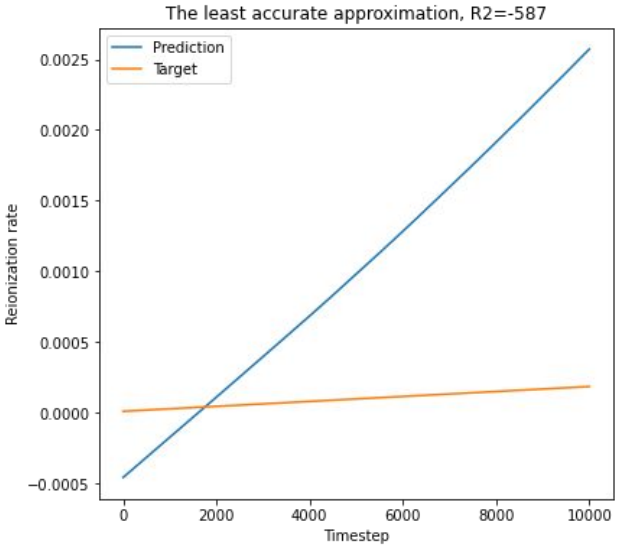


R2-score of best simulation: 0.99

R2-score of worst simulation: -15411534

After modifications



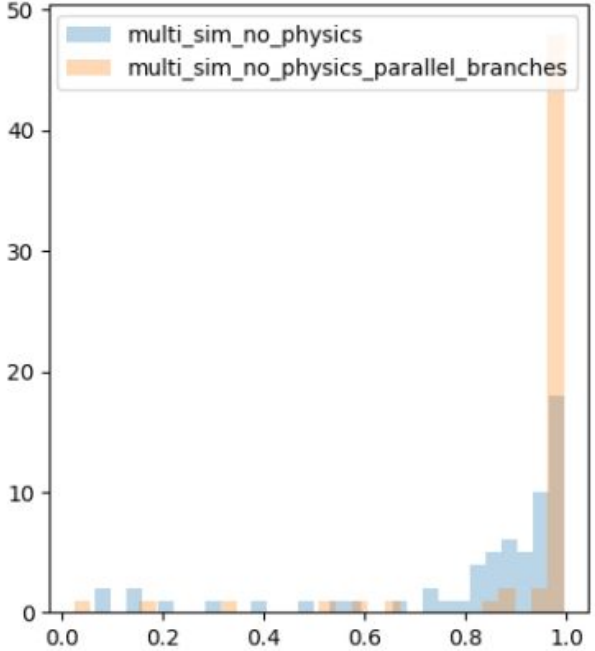


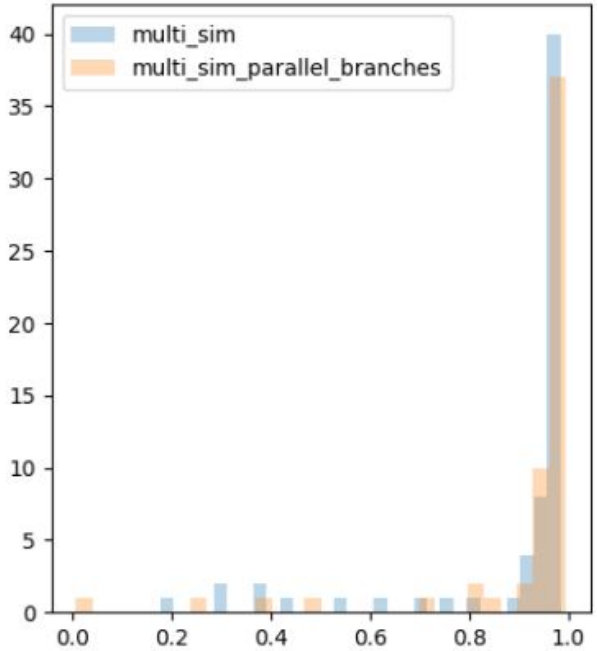
R2-score of best simulation: 0.98224294

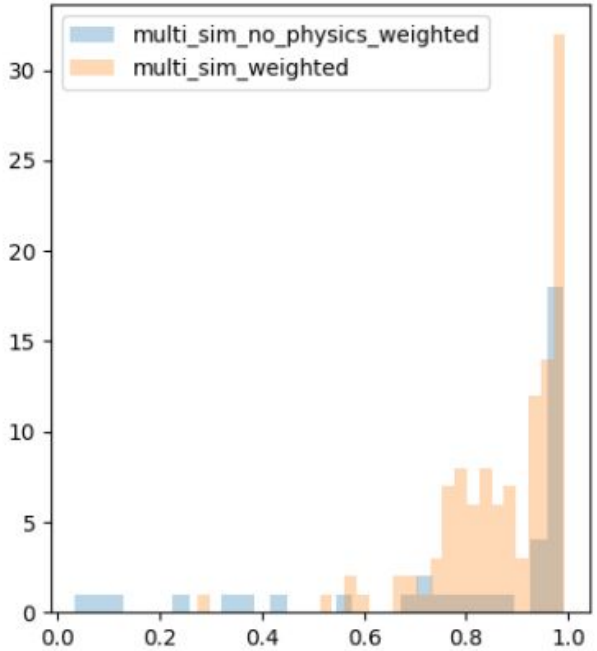
R2-score of worst simulation: -587.18427

Model is able to adapt for slowly evolving functions

Distributions of R2 scores for different configurations







Plan for next week

- finish writing the report
- finalize project