Dalle Molle Institute for Artificial Intelligence
IDSIA

A joint research institute of:
Università della Svizzera italiana
Scuola universitaria professionale della Svizzera italiana SUPSI

In-context learning for model-free system identification

M. Forgione, F. Pura, and D. Piga
IDSIA Dalle Molle Insitute for Artificial Intelligence
{marco.forgione, filippo.pura, dario.piga}@idsia.ch

Meta-learning of dynamical systems

In standard identification, we improve as we see more data from a single system. Can we also learn from similar ones?

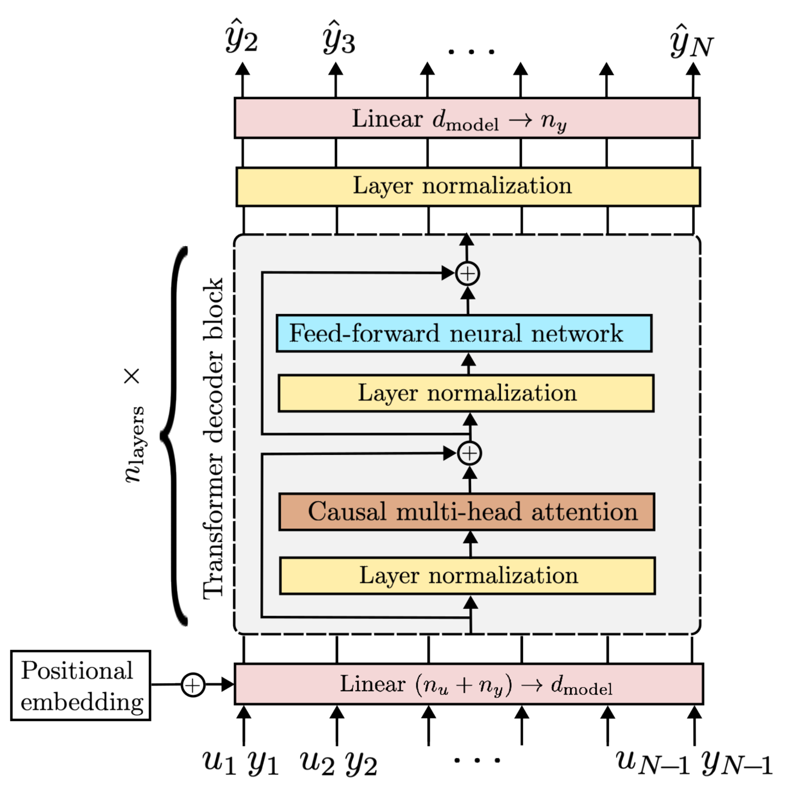
- Collection of datasets $\mathcal{D}^{(i)} = (u^{(i)}_{1:N}, y^{(i)}_{1:N})$ from different, but related dynamical systems $S^{(i)}$ available
- Can we get better at identifying $S^{(i)}$ as we observe more datasets $\mathcal{D}^{(j)}$? Can we learn to learn dynamical systems?

In-context learning approach

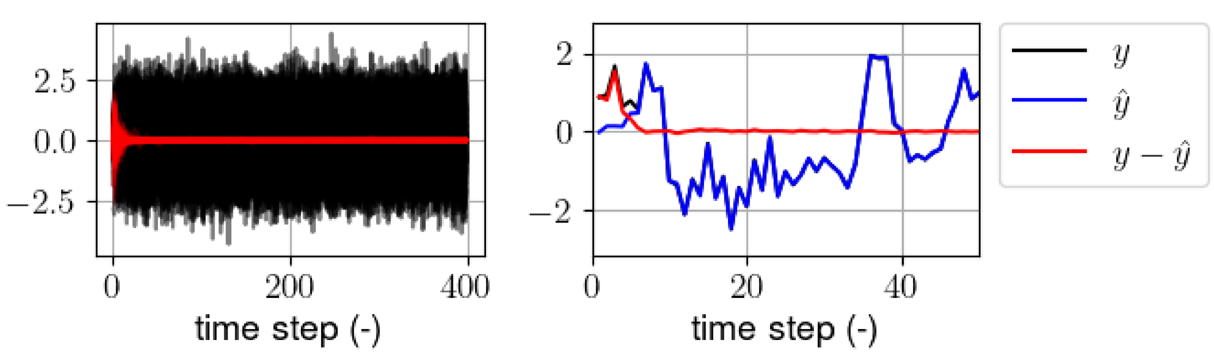
- Transformers are expressive as a programming language. We can train them to behave like algorithms
- We provide them with a context of input/output data and a task. They must learn to identify systems to solve the task
- If we manage to train the Transformer over a class of dynamical systems, it becomes a meta model of that class!

One-step prediction

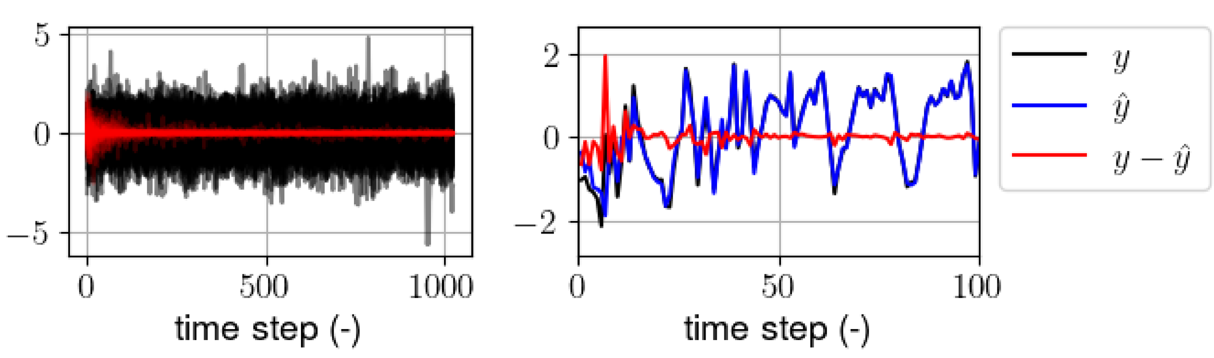
$$\hat{y}_{k+1} = \mathcal{M}_\phi(u_{1:k}, y_{1:k})$$



Linear system class

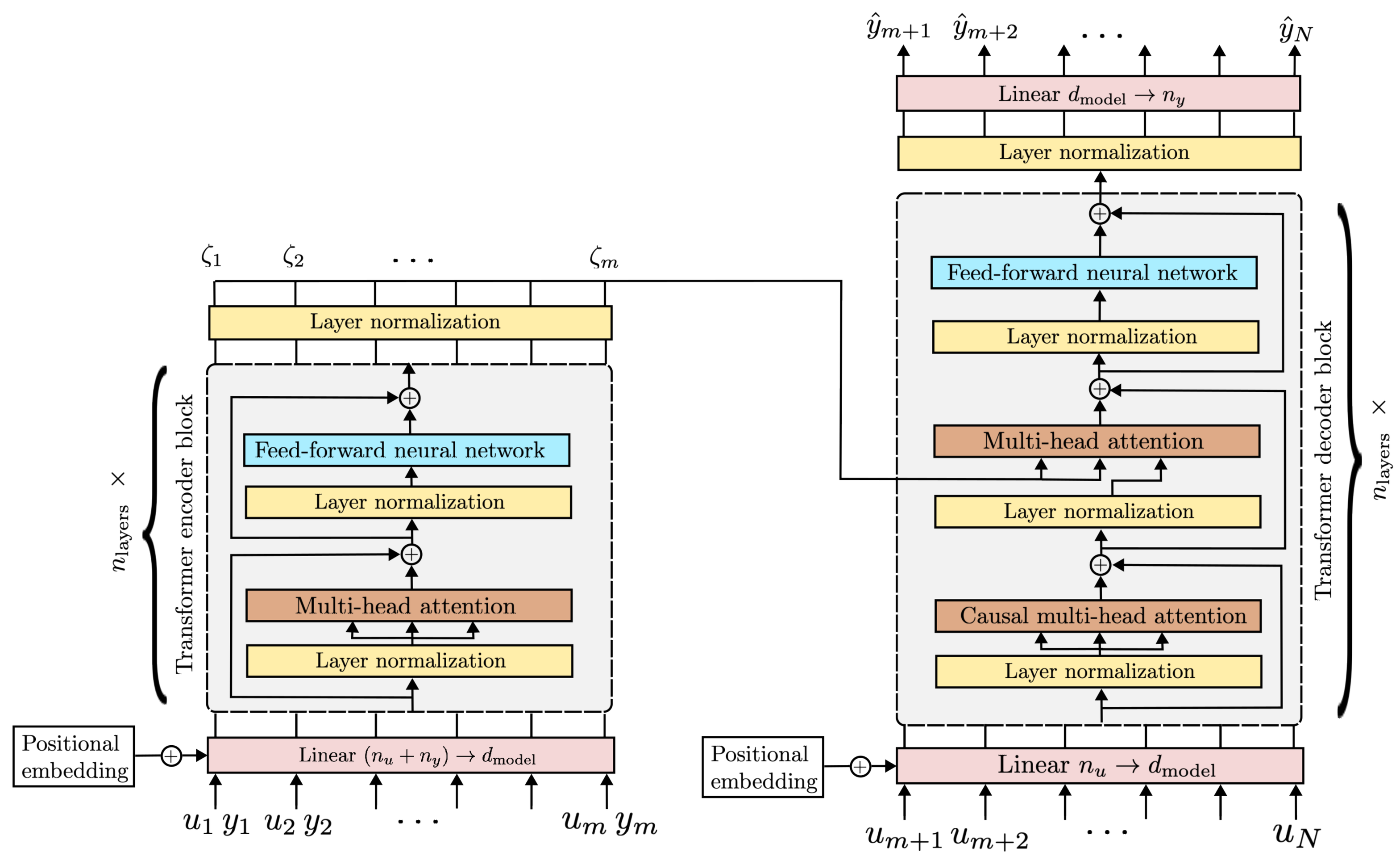


Wiener-Hammerstein system class

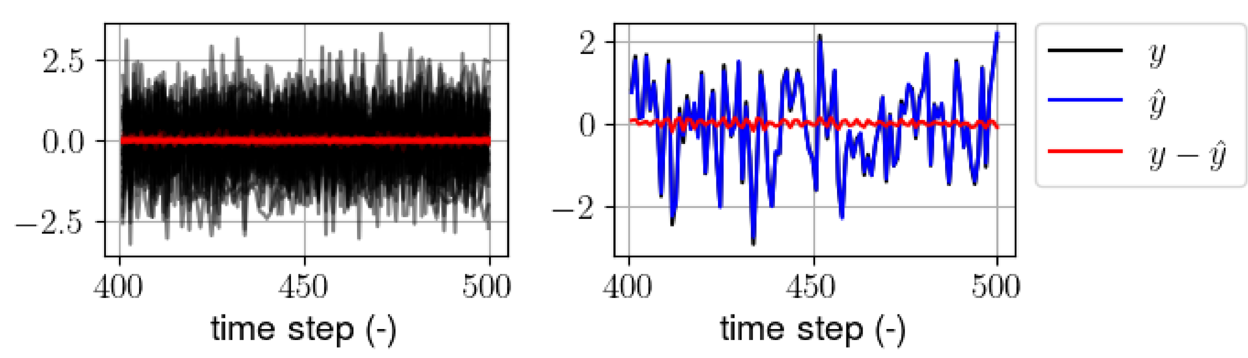


Multi-step simulation

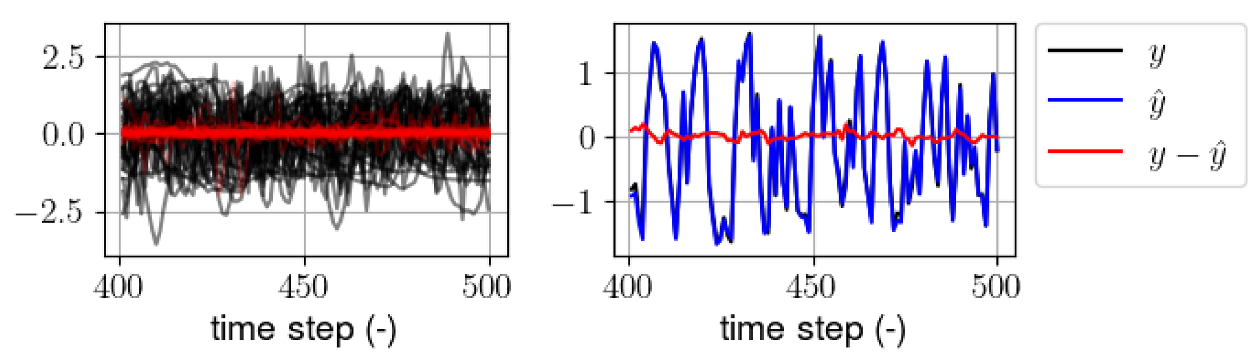
$$\hat{y}_{m+1:N} = \mathcal{M}_\phi(u_{1:m}, y_{1:m}, u_{m+1:N})$$



Linear system class



Wiener-Hammerstein system class



Future works

- Transfer learning from a system class to another one and fine tuning to a specific system instance
- Curriculum learning to solve complex tasks starting from models learned on simpler ones
- Analysis of the effect of noise during meta training. Can it help generalization? Does it hinder optimization?