OUR Project Efficient State Space Models through the Convolutional Lens

Spring 2025 Project Proposal

Project Title

Efficient State Space Models through the Convolutional Lens

Project Description

Recently, structured state-space models (SSMs) have emerged as a strong contender for sequence modeling in deep learning, mitigating semantic challenges of Transformer models induced by their limited window lengths. Recent such models have dominated certain benchmarks such as the Long Range Arena and have been especially successful in domains involving continuous signal data such as audio and vision, while also beating Transformers due to the linear or near-linear scaling of their computational complexity in sequence length. While the scaling of state-of-the-art SSMs such as Mamba and Mamba-2 is near-linear in the sequence length, it is quadratic in the hidden state dimension, which is detriment towards the construction of scalable, expressive SSMs capturing complex state information. In this project, we revisit a convolutional perspective on SSMs that additionally the leverage selection mechanism introduced by Mamba, and develop a variant of Mamba-2 type of model via a block-Toeplitz matrix perspective, which allows for a near-linear scaling in the hidden state dimension both for training and inference. Tasks in this project include the implementation and modification of SSM code bases, the training of SSM models, and the structuring and presentations of resulting computational experiments.

Number of Openings

Requirements

Students with a good grasp of calculus, linear algebra and strong interest in machine learning and good coding skills (Python) is a a requirement of this project. Some familiarity and lower-level understanding of the training process of deep learning models is also required.

Preferences

A familiarity with the fast Fourier transform and structured state-space models is helpful, but can be acquired during the project.

Training

No specific requirements, training takes places during the research project meetings and in discussions with the faculty and graduate student mentor.

Anticipated Student Learning Outcomes

- Familiarization with an important type of state-of-the-art machine learning model
- Experience with the development of new deep learning models
- Experience in contributing to ongoing AI/machine learning research
- If successful, co-authorship in a conference publication at a top-tier AI/ML conference can arise from this project. ## Mentoring Plan The student will be guided through the research process in weekly meetings with the mentoring faculty as well as in communication with a Ph.D. student, who serves as a co-mentor. Communication will occur additionally asynchronously through a Slack channel, during which smaller questions can be addressed. If the research project turns out to be successful, mentoring in academic writing and the preparation of research manuscripts will be provided.