

CURRENT PROJECTS:

- ***Neural Collapse Inspired Novel Metrics for Deciphering DNNs Classifiers and Their Applications:*** In this project, we undertake three interconnected tasks. First, we formulate metrics to assess the training quality of feature extractors and classifiers, and extensively evaluate them on a large set of DNN classifiers, along with their corresponding training and testing datasets. Next, using these metrics, we develop methods to evaluate pretrained DNNs with and without access to datasets, which are then used to select the most suitable DNNs for transfer and few-shot learning. Finally, we design methods that leverage these metrics to develop loss functions for training feature extractors, with the goal of reducing overfitting on the training data and narrowing the generalization gap.
- ***Mathematical Models for Predicting Performance of Image Segmentation Models:*** In this project, we engage in developing mathematical models that predict the performance of image segmentation models as a function of training dataset size and model architecture (with a fixed structure). We aim to empirically demonstrate that these models can estimate performance for any dataset size, compute asymptotic performance, and compare the relative effectiveness of different machine learning (ML) models. Additionally, we aim to demonstrate that they can be used to determine the cost-optimal training dataset size.
- ***Uncertainty Quantification of Hallucinations in LLMs:*** Building on recent advances in semantic entropy-based hallucination detection, this project focuses on developing a method that not only clusters outputs by semantic equivalence but also quantifies the epistemic uncertainty of semantic entropy itself. We leverage a quantum tensor network-based uncertainty quantification pipeline to enable more robust detection by modeling perturbation-induced uncertainty. Furthermore, we introduce an entropy-maximization-based selection strategy that prioritizes responses with high certainty and semantic coherence.