Learning Dynamics: the Key to Understanding Generalization and Memorization

Xiaolin Huang
Shanghai Jiao Tong University, China

For shallow models, generalization analysis primarily focuses on functional space properties, while training mainly influences computational efficiency. In contrast, deep models operate in an over-parameterized regime, where generalization performance is strongly tied to learning dynamics. This is because different global minima that achieve high training accuracy could exhibit varying performance on new data. Our research has demonstrated that learning dynamics can be effectively represented in a low-dimensional space, leading to improvements in both training efficiency and generalization performance. Studying these dynamics also provides valuable insights into how deep models memorize data. Beyond its theoretical significance, understanding memorization also plays a crucial role in machine unlearning — an emerging and impactful field with important real-world applications.